11-5-2019

Evolutionary Genetics of the Genus Zamia (Zamiaceae, Cycadales)

Michael Calonje
mcalo002@fiu.edu

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

EVOLUTIONARY GENETICS OF THE GENUS *ZAMIA* (ZAMIACEAE, CYCADALES)

A dissertation submitted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

in

BIOLOGY

by

Michael Calonje Bazar

2019

To: Dean Michael R. Heithaus
College of Arts, Sciences, and Education

This dissertation, written by Michael Calonje Bazar, and entitled Evolutionary Genetics of the Genus *Zamia* (Zamiaceae, Cycadales), having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

_____
Timothy M. Collins

_____
M. Patrick Griffith

_____
Hong Liu

_____
Alan W. Meerow

_____
Jennifer H. Richards

_____
Andrew P. Vovides

_____
Javier Francisco-Ortega, Major Professor

Date of Defense: November 5, 2019

The dissertation of Michael Calonje Bazar is approved.

_____
Dean Michael R. Heithaus
College of Arts, Sciences, and Education

_____
Andrés G. Gil
Vice President for Research and Economic Development
and Dean of the University Graduate School

Florida International University, 2019

DEDICATION

I dedicate this dissertation to my beloved wife Claudia and my daughters Alicia and Emma, the latter two born during the course of my studies at Florida International University. Without their patience, support, and understanding this dissertation would not have been possible.

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my major professor, mentor and friend Dr. Javier Francisco Ortega for his invaluable guidance and support during my academic journey at F.I.U. I would also like to extend my deepest gratitude to Dr. Alan Meerow for his expert mentorship in many different aspects of genetics and allowing access to the facilities, equipment and staff expertise at the USDA-ARS Subtropical Horticulture Research Station (Miami, FL) during our collaborative research. I am also grateful to others who also mentored and assisted me at the USDA-ARS laboratory, including Dayana Salas-Leiva, Kyoko Nakamura, and Vanessa Sánchez Vélez. I am extremely grateful to Montgomery Botanical Center (Miami, FL) and especially to Executive Director M. Patrick Griffith for his unconditional support in all aspects of my pursuit of this degree. I would like to extend my sincere thanks to Dr. Andrew Vovides (Instituto de Ecología, A.C., Xalapa, Veracruz, Mexico) for his expert mentorship in cycad anatomy and invaluable colleagueship, and to express my deepest appreciation to the rest of my committee (Dr. Timothy Collins, Dr. Jennifer Richards and Dr. Hong Liu) for their invaluable feedback and support. Finally, I am forever indebted to my beloved wife, Claudia Calonje for her unconditional support during these very busy times.

ABSTRACT OF THE DISSERTATION

EVOLUTIONARY GENETICS OF THE GENUS *ZAMIA* (ZAMIACEAE, CYCADALES)

by

Michael Calonje

Florida International University, 2019

Miami, Florida

Professor Javier Francisco-Ortega, Major Professor

The genus *Zamia* L. (Zamiaceae), consisting of 77 species, is the most species-rich and widely distributed cycad genus in the New World and is arguably the most morphologically and ecologically diverse genus in the Cycadales. We utilized a multilocus sequence dataset of 10 independent loci (9 single copy nuclear genes + 1 plastid) and extensive taxon sampling (over 90% of species) to infer phylogenetic relationships within *Zamia.* We infer a robust phylogenetic tree for the genus with a strong geographic delimitation of clades and find that four morphological characters typically used for diagnostic purposes in the genus exhibit a high degree of homoplasy.

We genotyped four populations of the Belizean endemic cycad *Zamia decumbens* using ten microsatellite loci and analyzed the data using a variety of population genetic analyses methods. The populations occurred in two different habitat types: inside dolines (one at a cave entrance and two at the bottom of sinkholes) and one on steep karstic terrain on a hilltop. We found the genetic variation was not structured geographically or by habitat type, but rather seemed to reflect the demographic history of the populations and their genetic connectivity. Contemporary geneflow between populations is generally low, with the Cave population being the most important population in facilitating genetic connectivity in the region, mostly as a source of migrants to other populations.

A conservation assessment for the three cycad species native to the Bahamas Islands is presented using field surveys on all islands where these species occur. *Zamia angustifolia* is native to Eleuthera, *Zamia integrifolia* is native to Abaco, Andros, Eleuthera, Grand Bahama and New Providence,

and *Zamia lucayana* is endemic to Long Island. We assessed the genetic structure of *Z. lucayana* using 15 polymorphic microsatellite DNA loci; this indicated that the three known populations should be considered a single management unit. However, the high number of private alleles suggests that genetic drift, indicative of recent fragmentation, is progressing. We propose in situ conservation strategies, and we also collected germplasm from a total of 24 populations of these three cycad species, for ex situ conservation.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

CHAPTER II

CHAPTER III

TABLE PAGE

LIST OF FIGURES

CHAPTER I

FIGURE PAGE

CHAPTER I (APPENDIX B)

FIGURE PAGE

CHAPTER I (APPENDIX C)

FIGURE PAGE

CHAPTER I (APPENDIX D)

FIGURE PAGE

CHAPTER I (APPENDIX E)

FIGURE PAGE

CHAPTER I (APPENDIX F)

FIGURE PAGE

CHAPTER II

FIGURE PAGE

CHAPTER II (APPENDIX A)

FIGURE PAGE

CHAPTER III

FIGURE PAGE

INTRODUCTION

Cycads, belonging to the order Cycadales, are long-lived, slow-growing, dioecious, pachycaul plants with starch-rich stems bearing terminal crowns of pinnate leaves. They are distinguished from other plants by a combination of morphological characters, including, a lack of axillary buds, the production of specialized roots hosting nitrogen-fixing cyanobacteria (coralloid roots), the presence of unique chemical compounds (e. g., cycasins, macrozamins and BMAA (the amino acid β-N-methylamino-L-alanine), and rudimentary ecological interactions with symbiotic cyanobacteria and insect pollinators (Norstog & Nichols, 1997). All cycad species are tropical or subtropical, most commonly occurring at low elevation. Populations are typically small, isolated and restricted to narrow geographic areas. They are considered the world's most threatened land plants (Marler & Marler, 2015, Fragnière *et al.*, 2015, Brummitt *et al.*, 2015) and are highly threatened by human activities, especially habitat destruction and the extraction of plants from the wild for horticultural (Donaldson, 2003) and traditional uses (Cousins, Williams & Witkowski, 2012). Cycads include 10 genera and a total of 355 currently accepted species (Calonje, Stevenson & Osborne, 2019b). The genus *Cycas* is the sole representative of the family Cycadaceae, whereas the other nine genera belong to the Zamiaceae (Christenhusz, 2011). Four endemic genera occur in the New World: *Ceratozamia*, *Dioon*, *Microcycas*, and *Zamia*. The genus *Zamia* L. (Zamiaceae, Cycadales), with 80 accepted species (Calonje *et al.*, 2019b), is the most species-rich and broadly distributed genus among the New World genera, and it is broadly considered the most ecologically and morphologically diverse genus of the extant cycads (Norstog & Nichols, 1997). Despite the relative importance of the genus within the cycadales, the phylogenetic relationships between species remain unclear, and only a very limited amount of research has been published into how their populations are structured. The focus of the present dissertation is to provide insight into the evolutionary history of the genus *Zamia* as a whole, as well as into the population genetics of two species belonging to two disjunct geographic regions (*Z. decumbens* and *Z. lucayana*).

The phylogenetic trees published prior to the present work that either focus on *Zamia* or contain a considerable sampling of the genus are of limited utility due to low taxon sampling (e.g., Clugston *et al.*, 2016) and / or a weak phylogenetic signal resulting from a small number of molecular markers (e.g.,

Caputo *et al.*, 2004, Nagalingum *et al.*, 2011). In Chapter I we utilize a large multilocus sequence dataset of 10 independent loci (9 SCNGs + 1 Plastid) and extensive taxon sampling (over 90% of species) to infer phylogenetic relationships within the genus *Zamia* using Maximum Parsimony, Maximum Likelihood, and Bayesian time-calibrated species-tree inference methods. We discuss our results in the context of clade age estimation, diversification, and biogeographical history.

In Chapter II we present a population genetics study of *Zamia decumbens*, a species endemic to the Maya Mountains of Belize, where it occurs in small disjunct populations over karst topography in two different habitat types. Some populations occur within dolines in areas protected from direct rainfall by overhangs, and others in the upper parts of ridges and hilltops, where soils are extremely rocky and well drained. We sampled three populations inside dolines (one at a shallow cave entrance and two at the bottom of deep, steep-walled sinkholes), and one on a rocky hilltop, utilizing 10 polymorphic SSR microsatellite DNA loci to examine the genetic diversity, genetic structure, and demographic history of these populations using a variety of analytical methods. We discussed the demographic history of this species as it relates to the phylogenetic history of the genus (Calonje *et al.*, 2019a) as well as the natural and geological history of the region. Finally, we discuss our findings under a conservation genetics framework.

In Chapter III, we surveyed *Zamia* populations throughout their entire known distribution range on six different islands of the Bahamas archipelago in order to provide information on their geographic distribution and their conservation status. This included: populations of *Z. integrifolia* on Abaco, Grand Bahama, New Providence and Eleuthera, *Z. angustifolia* on Eleuthera, and *Z. lucayana* on Long Island. For *Z. lucayana*, we assessed the genetic structure of *Z. lucayana* utilizing 15 polymorphic microsatellite DNA loci and discussed the conservation implications of our findings.

LIST OF REFERENCES

Brummitt, NA, Bachman, SP, Griffiths-Lee, J, Lutz, M, Moat, JF, Farjon, A, Donaldson, JS, Hilton-Taylor, C, Meagher, TR, Albuquerque, S, Aletrari, E, Andrews, AK, Atchison, G, Baloch, E, Barlozzini, B, Brunazzi, A, Carretero, J, Celesti, M, Chadburn, H, Cianfoni, E, Cockel, C, Coldwell, V, Concetti, B, Contu, S, Crook, V, Dyson, P, Gardiner, L, Ghanim, N, Greene, H, Groom, A, Harker, R, Hopkins, D, Khela, S, Lakeman-Fraser, P, Lindon, H, Lockwood, H, Loftus, C, Lombrici, D, Lopez-Poveda, L, Lyon, J, Malcolm-Tompkins, P, McGregor, K, Moreno, L, Murray, L, Nazar, K, Power, E, Quiton Tuijtelaars, M, Salter, R, Segrott, R, Thacker, H, Thomas, LJ, Tingvoll, S, Watkinson, G, Wojtaszekova, K, Nic Lughadha, EM. 2015. Green Plants in the Red: A Baseline Global Assessment for the IUCN Sampled Red List Index for Plants. *PLOS ONE* **10:** e0135152.

Calonje, M, Meerow, AW, Griffith, MP, Salas-Leiva, D, Vovides, AP, Coiro, M, Francisco-Ortega, J. 2019a. A Time-Calibrated Species Tree Phylogeny of the New World Cycad Genus *Zamia* L. (Zamiaceae, Cycadales). *International Journal of Plant Sciences*.

Calonje, M, Stevenson, DW, Osborne, R. 2019b. The World List of Cycads, online edition (http://cycadlist.org) Accessed (Accessed 21 Sep 2019).

Caputo, P, Cozzolino, S, Luca, PD, Moretti, A, Stevenson, DW. 2004. Molecular phylogeny of *Zamia* (Zamiaceae). In: Walters, TW, Osborne, R eds. *Cycad Classification: Concepts and Recommendations.* Wallingford, United Kingdom: CABI Publishing.

Christenhusz, MJMC, Mark W.;Fay, Michael F. 2011. Linear sequence, classification, synonymy, and bibliography of vascular plants: Lycophytes, ferns, gymnosperms and angiosperms. *Phytotaxa* **19:** 1-134.

Clugston, JAR, Griffith, MP, Kenicer, GJ, Husby, CE, Calonje, MA, Stevenson, DW, Little, DP. 2016. *Zamia* (Zamiaceae) phenology in a phylogenetic context: does in situ reproductive timing correlate with ancestry? *Edinburgh Journal of Botany* **73:** 345-370.

Cousins, SR, Williams, VL, Witkowski, ETF. 2012. Uncovering the cycad taxa (Encephalartos species) traded for traditional medicine in Johannesburg and Durban, South Africa. *South African Journal of Botany* **78:** 129-138.

Donaldson, JS. 2003. Introduction. In: Donaldson, JS ed *Cycads: Status, Survey, and Conservation Action Plan.* Gland, Switzerland: IUCN/SSC Cycad Specialist Group. IUCN.

Fragnière, Y, Bétrisey, S, Cardinaux, L, Stoffel, M, Kozlowski, G. 2015. Fighting their last stand? A global analysis of the distribution and conservation status of gymnosperms. *Journal of Biogeography* **42:** 809-820.

Marler, PN, Marler, TE. 2015. An Assessment of Red List Data for the Cycadales. *Tropical Conservation Science* **8:** 1114-1125.

Nagalingum, NS, Marshall, CR, Quental, TB, Rai, HS, Little, DP, Mathews, S. 2011. Recent Synchronous Radiation of a Living Fossil. *Science* **334:** 796-799.

Norstog, KJ, Nichols, TJ 1997. *The biology of the cycads* Ithaca, New York: Cornell University Press.

CHAPTER I

A time-calibrated species tree phylogeny of the New World cycad genus *Zamia* L. (Zamiaceae, Cycadales)

MICHAEL CALONJE BAZAR[a,b*], ALAN W. MEEROW[c], M. PATRICK GRIFFITH[b], DAYANA SALAS-LEIVA[a,b,c,g], ANDREW P. VOVIDES[d], MARIO COIRO[e], and JAVIER FRANCISCO-ORTEGA[a, f]

[a]Department of Biological Sciences, Florida International University, Miami, FL 33199, U.S.A.; [b]Montgomery Botanical Center, Coral Gables, FL 33156, U.S.A.; [c]USDA-ARS-SHRS, Miami, FL 33158, USA; [d]Instituto de Ecología, A.C., Xalapa, Veracruz 91000, Mexico; [e]Department of Systematic and Evolutionary Botany, University of Zurich, Switzerland; [f]Fairchild Tropical Botanic Garden, Coral Gables, FL 33156, USA; [g]Centre for Comparative Genomics and Evolutionary Bioinformatics (CGEB), Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, NS, Canada, B3H 4R2.

[*] For correspondence: michaelc@montgomerybotanical.org

A modified version of Chapter 1 was published in International Journal of Plant Sciences 180(4):286–314. 2019; with the same title and authors.

ABSTRACT

Premise of research. The genus *Zamia* L. (Zamiaceae), consisting of 77 species, is the most species-rich and widely distributed cycad genus in the New World and is arguably the most morphologically and ecologically diverse genus in the Cycadales. However, a strong phylogenetic framework for this genus is still lacking.

Methodology. We utilized a multilocus sequence dataset of 10 independent loci (9 single copy nuclear genes + 1 plastid) and extensive taxon sampling (over 90% of species) to infer phylogenetic relationships within *Zamia.* We implemented concatenated matrix analyses with maximum parsimony and maximum likelihood methods as well as a time-calibrated Bayesian species tree-estimation approach. Diversification, historical biogeography, and ancestral character state reconstruction analyses were conducted utilizing the species tree topology which was the most morphologically and geographically congruent.

Pivotal results. We infer a robust phylogenetic tree for the genus with a strong geographic delimitation of clades and find that four morphological characters typically used for diagnostic purposes in the genus exhibit a high degree of homoplasy. The stem group of *Zamia* was estimated to have originated at 68.28 Ma (95% HPD 51.0–84.5), and the crown group at 9.54 Ma (95% HPD 9.0–10.62). The majority of species richness in the genus appeared during the Pliocene and Pleistocene, with the highest diversification rates occurring in clades comprised of Caribbean and South American species. Biogeographic analysis suggests a Caribbean or Mesoamerican origin for *Zamia* with subsequent dispersal to the Central American Isthmus and South America, where the genus reaches its maximum species and morphological diversity.

Conclusions. The high degree of convergence found in multiple morphological characters has historically confounded efforts to elucidate species relationships utilizing non-phylogenetic methods. This study presents the most species-comprehensive, well-resolved hypothesis of phylogenetic relationships within *Zamia* and should provide a strong phylogenetic framework for further studies.

INTRODUCTION

The genus *Zamia* L. (Zamiaceae, Cycadales) is widely considered to be the most ecologically and morphologically diverse genus of the extant cycads (Norstog & Nichols, 1997). With its 77 accepted species (Calonje, Stevenson & Osborne, 2019), it is the most species-rich and broadly distributed genus among the New World genera of the Zamiaceae. It is restricted primarily to the Neotropical region (sensu Morrone, 2014, Sclater, 1858) with only the northernmost *Zamia integrifolia* L.f. populations in Florida and southeast Georgia (U.S.A.) extending into the Nearctic region (sensu Escalante, Morrone & Rodríguez-Tapia, 2013, Morrone, 2014). The distribution of the genus can be spatially divided into three separate areas of endemism: 1) a Caribbean group restricted to islands on the Bahama Archipelago and the Greater Antilles as well as to the mainland states of Florida and Georgia in the Southeastern U.S.A.; 2) a Mesoamerican group extending from Tamaulipas, Mexico, to northern El Salvador; and 3) a Central and South American group occurring from southern Nicaragua to Bolivia, the southernmost extent of the genus (fig. 1).

Despite the remarkable species diversity and broad geographic distribution of *Zamia*, the only formal infrageneric classification of the genus was published decades ago in Schuster's (1932) monograph of cycads. Schuster included 26 species in his treatment of *Zamia* and divided the genus into three geographically defined Sections ('Caribaeae', 'Mexicano-Meridionales', and 'Centrali-Meridionales') which are broadly congruent with the Caribbean, Mesoamerican, and Central-and-South American species groups outlined above. However, Schuster did not designate types for these sectional names which remain untypified to this day, so their taxonomic application remains problematic. Moreover, Schuster did not follow nomenclatural rules requiring the subdivisions of genera (art. 22.1, McNeill *et al.*, 2012), making several of his sectional names, such as 'Caribaeae' illegitimate. Schuster's work is also notorious for contravening the rules of priority and other nomenclatural rules, and it includes very elaborate and often nonsensical infraspecific hierarchies (see Johnson, 1959, Hill, 1996). Consequently, Schuster's classification has been disregarded by most authors, although some of Schuster's sectional names have been typified and adopted in formal classification work, most notably in *Cycas* (see Hill, 1995). However, no authors since Schuster have attempted formal classification of supraspecific taxa within *Zamia.* Rather,

subsequent authors have mostly limited themselves to informally classifying and surmising interspecific relationships within subsets of the genus based on similarities in geographic distribution (Norstog & Nichols, 1997), morphology (e.g. Calonje *et al.*, 2010, Schutzman, Vovides & Dehgan, 1988, Calonje *et al.*, 2011, Taylor, Haynes & Holzman, 2008), anatomy (Acuña-Castillo & Marín-Méndez, 2013), karyology (Caputo *et al.*, 1996), genome sizes (Zonneveld & Lindström, 2016), and other characters. However, the phylogenetic relationships within the genus remain unclear, as only a limited number of studies have involved a species-level phylogenetic component.

Most notably, Caputo *et al.* (2004) conducted a phylogenetic analysis of *Zamia*, including 23 of the 77 currently accepted species in a parsimony-based phylogenetic analysis of sequences of the Internal Transcribed Spacer 2 (ITS2) of nuclear ribosomal DNA in combination with a morphological dataset (Fig. 2A). The analysis found several clades to be more congruent with geographic distribution than with morphological similarities, leading the authors to suggest that convergent evolution of morphological characters is pervasive in the genus. Geographically congruent clades identified included those comprised of Central American, North American, and Caribbean species, as well as a combined clade including South American and Central American species.

Clugston *et al.* (2016) examined the phenological phases of 11 species of *Zamia* in a phylogenetic context, presenting a tree based on a maximum parsimony analysis using nucleotide sequence data from two plastid genes (*matK*, *rbcL*), two single-copy nuclear genes (*CAB*, *NEEDLY*), and one high-copy nuclear gene (*26S*). This analysis showed strong support for a Caribbean clade and a monophyletic South American group, and moderate support for a Central American lineage (Fig. 2B).

To examine the diversification ages of extant cycads, Nagalingum *et al.* (2011) used the nuclear gene Phytochrome P (*PHYP*) with a few additional sequences from the chloroplast genes *matK* and *rbcL* to produce a time-calibrated phylogeny of 199 taxa, including 31 species of *Zamia*. The results indicated that most living cycad species are the result of quite recent diversification of a very ancient lineage and are therefore much younger than previously thought. This was previously suggested by Treutlein and Wink (2002) in an earlier phylogenetic study using *rbc*L plastid sequences and was also found to apply generally

to extant gymnosperms by Crisp and Cook (2011). Although infrageneric relationships were not discussed by Nagalingum *et al.* (2011), their phylogeny supported the monophyly of *Zamia*, albeit with low node support throughout most of the genus (Fig. 2C). Similarly, in a paper using the same molecular markers but with the inclusion of additional species and fossil calibration constraints, Condamine *et al.* (2015) also found poor node support within genera. Although both divergence time estimation studies included a large number of cycad species in their analyses, they similarly obtained generally low node support within genera, likely due to the sparse and often non-overlapping locus sampling between species and the limited signal present in the markers selected.

As illustrated above, the phylogenies published to date that focus on *Zamia* or contain a considerable sampling of the genus are of limited utility due to low taxon sampling (e.g. Clugston *et al.*, 2016) and / or a weak phylogenetic signal resulting from a small number of molecular markers (e.g. Caputo *et al.*, 2004, Nagalingum *et al.*, 2011). In this paper we utilize a larger multilocus sequence dataset of 10 independent loci (9 SCNGs + 1 Plastid) and extensive taxon sampling (over 90% of species) to infer phylogenetic relationships within the genus *Zamia* using Maximum Parsimony, Maximum Likelihood, and Bayesian time-calibrated species-tree inference methods. We discuss our results in the context of clade age estimation, diversification, and biogeographical history.

MATERIALS AND METHODS

TAXONOMIC SAMPLING

DNA was isolated from 113 individual leaflet samples of *Zamia* belonging to 70 of the 76 currently accepted species in the genus (Calonje *et al.*, 2019) and from single samples of *Microcycas calocoma* (Miq.) A.DC. and *Stangeria eriopus* (Kunze) Baill. that were included as outgroups for phylogenetic analyses. Multiple samples were collected for some species of *Zamia*, particularly from those with broader geographic distributions. The leaflet samples were obtained primarily from plants of known provenance or pedigree cultivated in botanic gardens and other botanical collections (Appendix A: Accessions sampled for molecular analyses).

OUTGROUP SELECTION

Microcycas calocoma and *Stangeria eriopus,* both belonging to monotypic genera, were selected as outgroups for phylogenetic analyses, the latter used as the functional outgroup for all phylogenetic analyses. *Microcycas* is strongly supported as the sister genus to *Zamia* and *Stangeria* as sister to *Zamia* + *Microcycas* in a recent genus-level phylogeny of the Cycadales that used multiple phylogenetic methods (Salas-Leiva *et al.*, 2013). The same phylogenetic arrangement of these three genera was also supported in other phylogenetic studies (e.g. Rai *et al.*, 2003, Chaw *et al.*, 2005, Zgurski *et al.*, 2008, Bogler & Francisco-Ortega, 2004).

MARKER SELECTION

Sequences of nine independent, cycad-specific single-copy nuclear genes (SCNG) and one chloroplast gene (table 1) were used to infer the phylogeny of the genus *Zamia* using various phylogenetic methods. Seven of these were identified and developed in our laboratories (Salas-Leiva *et al.*, 2014) and used in this study (*40S*, *ATG2*, *GroES*, *HTS*, *LiSH*, *PEX4*, *PMP22*, *WRKY4*); the remaining ones were from previously published SCNG (*CyAG*) and plastid (*psbK/I*) markers (table 2) that appeared informative in our *Zamia* test panel. The loci *40S*, *ATG2*, *CyAG* and *GroES* were previously used to infer the generic relationships of the Cycadales (Salas-Leiva *et al.*, 2013), and *40S*, *ATG2*, *CyAG*, *GroES*, *LiSH*, *PEX4* and *WRKY4* were applied to assess the genetic diversity and genetic structure of Bahamian zamias (Salas-Leiva *et al.*, 2017). The single chloroplast gene locus included in the study (*psbK/I*) was selected because it had displayed the best performance (amplification success and high number of diagnostic sites) in a previous work that evaluated seven different chloroplast genes for barcoding Mexican species of *Zamia* (Nicolalde-Morejón *et al.*, 2011).

DNA EXTRACTION, AMPLIFICATION, AND SEQUENCING

DNA was extracted from 20-100mg of fresh or silica gel-desiccated leaflet tissue using the FastDNA® SPIN Kit (MP Biomedicals, Santa Ana, CA, USA) and the FastPrep® -24 Instrument (Qbiogene, Inc., CA). The extracted DNA was quantified using GeneQuant Pro RNA/DNA Calculator Spectrophotometer (GE Healthcare Life Sciences) or NanoDrop 2000 Spectrophotometer (Thermo Scientific) and diluted to 10 ng/

µL. Amplifications contained 1X amplification Buffer with 2 mM $MgSO_4$, 10 mM of dNTPs, 0.2 mg/mL bovine serum albumin, 10 µM of forward and reverse primer, 0.05 U/µL Taq DNA polymerase (New England BioLabs, Ipswich, MA, USA) and 10 ng/µL of template DNA brought to a total volume of 15 µL with nuclease-free H_20. DNA samples were amplified using C1000 and S1000 thermal cyclers (Bio-Rad Laboratories, Hercules, CA). Most genes were amplified in single fragments, with the exception of *CyAG* and *GroES* that were each amplified in two separate fragments using separate primer pairs. The following thermal profile was used for all loci except *GroES*: 95°C for 2 min, 35 cycles of 95°C for 30 s, annealing temperature (50–60 °C) for 1 min, 72 °C for 1 min, and a final extension at 72°C for 7 min. Both fragments of the locus *GroES* were amplified using touchdown PCR (95°C for 2 min; three cycles of 95°C for 30 s, 55°C for 1 min, 72°C for 1 min; three cycles of 95°C for 30 s, 54°C for 1 min, 72°C for 1 min; three cycles of 95°C for 30 s, 53°C for 1 min, 72°C for 1 min; three cycles of 95°C for 30 s, 52°C for 1 min, 72°C for 1 min; three cycles of 95°C for 30 s, 52°C for 1 min, 72°C for 1 min; three cycles of 95°C for 30 s, 51°C for 1 min, 72°C for 1 min; 30 cycles of 95°C for 30 s, 50°C for 1 min, 72°C for 1 min; and final extension of 72°C for 10 min).

The PCR amplifications were evaluated by electrophoresis using 1.2% agarose gel stained with GelRed (Biotium, Inc., Hayward, CA, USA) and a size standard ladder (100 bp New England Biolabs). PCR products were purified using exonuclease I (New England Biolabs) and shrimp alkaline phosphatase (USB Products- Affymetrix, Santa Clara, CA, USA), incubating at 37°C for 1 h, followed by 80°C for 20 min. Single sequencing reactions used 1-2 µL of purified PCR product and were performed using ABI Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Carlsbad, CA, USA), followed by ethanol clean up. Labeled fragments were visualized on an ABI 3730 Automatic DNA Sequencer (Applied Biosystems), and the nucleotide sequences were manually edited with Sequencher 4.9 (Gene Codes Corporation, Ann Arbor, MI, USA). Sequence data was obtained for all samples across all loci except for *Stangeria eriopus*, which failed to amplify with the PCR markers used for the 40S and LiSH loci. All DNA sequence data (1148 sequences) in this study were deposited in GenBank (Appendix A: Accessions sampled for molecular analyses)

ALIGNMENT

Sequences for each locus were aligned using MAFFT version 7 (Katoh & Standley, 2013) and/or aligned manually with Sequencher 4.9. The aligned lengths ranged from 498 bp (*PEX4*) to 1991 bp (*GroES*). Sequence data for different loci were joined for concatenated analyses using Mesquite 3.2 (Maddison & Maddison, 2017).

The *PSBK/I* chloroplast locus exhibited a 180 bp inversion in the sequences of 16 samples encompassing seven species occurring in the Central American Isthmus (*Zamia cunaria* Dressler & D.W.Stev., *Z. dressleri* D.W.Stev., Z. elegantissima Schutzman, Vovides & R.S.Adams, *Z. imperialis* A.S.Taylor, J.L.Haynes & Holzman, *Z. nana* A.Lindstr., Calonje, D.W.Stev. & A.S.Taylor, *Z. obliqua* A.Braun, *Z. pseudoparasitica* J.Yates, and *Z. stevensonii* A.S.Taylor & Holzman), necessitating the manual realignment of the inverted regions. The inversions could be dealt with by either reverse-complementing the inverted sequence region, or inserting an 180 base pair gap in the alignment for all sequences that lacked the inversion (Morrison, 2009). We chose the latter as it was not clear whether the single nucleotide polymorphisms observed resulted from mutations prior to or after the inversion event.

PHYLOGENETIC ANALYSES

Concatenated matrix analyses

Aligned, concatenated sequences for the ten loci were analyzed using Maximum parsimony (MP) and Maximum likelihood (ML) methods. Maximum parsimony analyses were performed in PAUP v. 4.10b (Swofford, 2003) for each gene separately (results not shown) and for the concatenated matrix. Heuristic searches were conducted using 1000 random stepwise-addition replicates, MULTTREES on, and saving up to 10 minimum length trees per search for swapping using tree bisection-reconnection (TBR) branch swapping. Jackknife analysis was conducted for the concatenated matrix with 37% deletion probability for each character (JK; Farris et al., 1996; 1000 replicates with simple stepwise-addition, TBR branch-swapping, saving no more than 10 trees per replicate, and retaining only groups with frequency > 50%). A non-partitioned Maximum Likelihood (ML) analysis was conducted using RAxML 8.0.24 (Stamatakis,

2014) as implemented in the CIPRES Science Gateway (Miller, Pfeiffer & Schwartz, 2011) using the GTR Gamma model and 1000 bootstrap re-samplings.

Time-calibrated species tree

A time-calibrated species tree analysis was performed using the multispecies coalescent model of *BEAST (Heled & Drummond, 2010) as implemented in BEAST (ver. 2.4.4, Bouckaert *et al.*, 2014). Samples were assigned to species according to individual taxonomic assessments, resulting in a species tree topology with 77 terminals.

Separate, uncorrelated, log-normal, relaxed-clock models (UCLD) and random starting trees were assigned to each partition and a birth death process was applied as branching process prior (or tree prior), as it has shown to be a better fit for cycads than the traditionally-used Yule process (Condamine *et al.*, 2015). The following priors were set following Condamine *et al.* (2015): a uniform prior between 0 and 10 with a starting value at 0.1 for the mean growth rate, a uniform prior between 0 and 1 with a starting value of 0.5 for the relative death rate, an exponential prior with a mean of 0.22 on the standard deviation of the UCLD model, and a uniform prior between 0 and 1 on the mean of the UCLD model. Site models were estimated independently for each partition (table 3) using a Bayesian approach as implemented by the bModelTest package (Bouckaert & Drummond, 2017) available in BEAST.

The species tree was calibrated using 95% highest posterior density (HPD) age estimates obtained from a previously published time divergence analysis of the Cycadales based on 6 fossil calibrations and the birth-death tree prior (Condamine *et al.*, 2015), as well as from a recently discovered *Zamia* fossil (Erdei *et al.*, 2018). Prior distributions for all calibrated nodes were conservatively set to uniform using the minimum and maximum age bounds outlined below.

The age intervals for the tree root node (74.3 to 147.7 Ma) and the crown node of *Zamia* (9 to 22.1 Ma), as well as the maximum age bound (84.5 Ma) for the stem node of *Zamia* were set according to the 95% HPD estimates provided by Condamine *et al.* (2015). The minimum age bound for the stem group of *Zamia* was constrained to the minimum age estimate (33 Ma) of a fossil cycad leaflet of unknown terrestrial origin recently found preserved in marine sediments of the Gatuncillo Formation in Central Panama, which is assignable to *Zamia* based on cuticular micromorphology (Erdei *et al.*, 2018). The fossil

age is considered to be 35-33 Ma based on nannoplankton and foraminiferal biostratigraphy, and the minimum age estimate (33 Ma) was used to constrain the minimum age bound for the stem group of *Zamia*. This age is comparable to the lower bound of the 95% HPD estimate (34.2 Ma) obtained by Condamine *et al.* (2015).

Two independent runs of 750 million Markov chain Monte Carlo (MCMC) iterations were conducted, sampling every 2,500 iterations and preceded by a 100 million iteration burn-in period. Log and tree files were combined using Logcombiner (ver. 2.4.4, included in BEAST package), resulting in 600,000 trees. The log output was evaluated using Tracer (ver. 1.6, Rambaut *et al.*, 2014). A Maximum Clade Credibility (MCC) tree was created from these trees using TreeAnnotator (ver. 2.4.4, included in BEAST package) and visualized using Figtree (ver. 1.4.3, Rambaut, 2012). The MCC tree topology recovered was used for the diversification, biogeographic, and ancestral character state reconstruction analyses presented here.

DIVERSIFICATION ANALYSES

Diversification rates were estimated per million-year interval spanning 10–0 Ma for the crown node of *Zamia* as well as several major clades within the genus using the dates derived from the time-calibrated species tree. The rates were obtained using Foote's (Foote, 2000) equation for per-capita origination rates as modified to apply to molecular tree lineages by Nagalingum *et al.* (2011). We examined changes in diversification rates for each of the major clades using the APE package for R (Paradis, Claude & Strimmer, 2004) to calculate the γ statistic (Pybus & Harvey, 2000) from the branching times in the time-calibrated species tree. The γ statistic is a metric of the distribution of speciation times in molecular phylogenies, with negative values indicating early diversification and positive values indicating late branching. A Monte Carlo constant rates test (MCCR) was conducted in order to determine whether the distribution of diversification rates significantly differed from a constant rate ($p < 0.05$). The MCCR test simulates a distribution for the γ statistic by simulating trees that take into account the incomplete lineage sampling of the observed tree and providing a critical value for rejecting a constant diversification rate. The MCCR test was conducted using the LASER package in R (Rabosky, 2006) by simulating 10,000 trees under a pure birth model. Lineage through time (LTT) plots were prepared using the time-calibrated

species tree using the APE package in R. We examined the stem and crown nodes of *Zamia* as well as several major clades within the genus.

Net diversification rates were calculated according to Magallon and Sanderson (2001) using the package Geiger (Harmon *et al.*, 2008), as implemented in R. The method requires the input of the extinction rate as a fraction of the speciation rate (e), and since the extinction rate cannot be reliably estimated, we followed previous recommendations (Magallon & Sanderson, 2001, Crisp & Cook, 2011) and provided net diversification rate estimates based on several values ranging from low (e=0) to high (e=0.9).

The Comet model (May, Höhna & Moore, 2016) from the R package TESS (Höhna, May & Moore, 2016) was used to test for shifts in speciation or extinction rates, as well as for the signature of mass extinction in the crown node of *Zamia*. We selected a single prior for the diversification shift to represent a slowdown after the origin of the crown of *Zamia*. For the mass extinction priors, we have chosen a prior with three events, representing (1) the first appearance of polar ice caps during the Late Miocene ca. 7 Mya; (2) the Messinian glaciation (6.26 to 5.5 Mya) and (3) the start of the Pleistocene glaciations (2.58 Mya) (Herbert *et al.*, 2016, Hodell *et al.*, 2001). The survival rate prior was set to 0.05, representing a major extinction event.

We used BAMM v. 2.5.0 (Rabosky, 2014) to test evolutionary rate heterogeneity in the topology of the crown node of *Zamia* using the default priors and settings recommended by the developers.

ANCESTRAL CHARACTER STATE RECONSTRUCTION

We used Mesquite version 3.02 (Maddison & Maddison, 2017) to study the evolution of four morphological characters on the calibrated species tree by performing ancestral character state reconstruction using parsimony. The morphological characters used were presence of prickles on the petiole, prominence of teeth along leaflet margin, prominence of leaflet veins, and arborescence (table 4). These characters were selected because they are commonly used to diagnose different species within the genus.

BIOGEOGRAPHIC ANALYSES

The historical biogeography of *Zamia* was inferred using a statistical dispersal-vicariance analysis (S-DIVA, Yu, Harris & He, 2010) as implemented in RASP v. 3.2 (Yu *et al.*, 2015). The MCC tree resulting from the time-calibrated species tree analysis was selected as the condensed tree, and the 600,000 individual post-burnin trees obtained from the analysis were re-sampled to a more manageable 14,000 trees using Logcombiner v 2.4.4. One hundred of these were randomly selected in RASP for the analysis. Areas coded for the analysis include A= Caribbean, B= Mesoamerica; C= Isthmus, D= South America, and E = South Africa. The maximum number of ancestral areas at each node was constrained to four, and the estimated probabilities of the ancestral areas were visualized on the condensed tree.

We examined the considerable gap in the geographic distribution of *Zamia* occurring between northern El Salvador / Honduras and southern Nicaragua, referred to here as the Mainland *Zamia* distribution gap, by utilizing MaxEnt software 3.4.1 (Phillips, Anderson & Schapire, 2006) with Chelsa land surface climate data set layers (Karger *et al.*, 2016) to model the potential distribution of the two major clades surrounding this gap (i.e. Mesoamerica and Isthmus clades), as well as of *Zamia soconuscensis* Schutzman, Vovides & Dehgan. We used 1251 presence records from a database of specimen and field observations derived from several years of *Zamia* systematics research by the first author, and secondarily from occurrence records downloaded from the Global Biodiversity Information Facility (GBIF.org, 2017). The latter were further processed by correcting nomenclatural and georeferencing errors. To deal with spatial autocorrelation issues derived from spatial clusters of observations, we used the R package spThin (Aiello-Lammens *et al.*, 2015) to spatially thin the occurrence records to a minimum nearest neighbor distance of 1 km, resulting in a total of 767 records after thinning.

MaxEnt analyses were performed using the default feature types (linear, quadratic, product, threshold, and hinge) with the 'auto features' option, and the default output format of cloglog, which provides an estimate between 0 and 1 for probability of presence. Model performance was evaluated using the area under the receiver operating curve (AUC).

We examined the evolution of genome size in *Zamia* by using the function contMap from the phytools R package (Revell, 2012) to map the haploid genome sizes (1C) compiled by Zonneveld and

Lindström (2016) onto a pruned phylogeny for the crown group of *Zamia*. We used the used the function phylosig to compute the phylogenetic signal using Pagel's lambda.

RESULTS

SINGLE LOCUS AND CONCATENATED MATRIX SEQUENCE CHARACTERISTICS

Single locus MP analyses generally yielded poorly resolved consensus trees (figs. B3-B12, available online), indicating that the incongruence between them is most likely "soft" incongruence due to the weak phylogenetic signal of most individual gene trees, rather than "hard" incongruence, which is due to significantly conflicting topologies (Wendel & Doyle, 1998). The range of parsimony informative characters (table 5) ranged from 29 (*PEX4*) to 127 (*GroES*) or by percentage, from 5.1% (*LiSH*) to 10.2% (*ATG2*). Consistency indices (CI) were above 0.82 and retention indices (RI) were above 0.89 for all loci examined topologies (Wendel & Doyle, 1998). The concatenated matrix consisted of 9998 characters, 7% of which were parsimony informative, and the MP analysis yielded a total of 1,820 equally parsimonious trees (tree length = 2776, CI = 0.728, homoplasy index (HI) = 0.272 and RI = 0.869).

MAJOR CLADES RECOVERED

The maximum likelihood (ML), maximum parsimony (MP) concatenated analyses, as well as the Bayesian inference (BI) species tree approach recovered the genus *Zamia* as monophyletic and sister to *Microcycas*. All analyses recovered the same broad topology consisting of the following strongly geographically delimited major clades (fig. 3): (1) Caribbean clade, consisting mostly of Caribbean Island species (a single species, *Z. integrifolia*, also reaches Florida) that is sister to the rest of the genus consisting of species occurring primarily in the mainland of the Americas; (2) Fischeri clade, consisting of three Mexican endemic species and itself sister to the rest of the genus excluding the Caribbean clade; (3) Mesoamerica clade, including all other species occurring in Mesoamerica to the exclusion of the Fischeri clade and *Z. soconuscensis*; (4) Isthmus clade, consisting primarily of Panamanian and Costa Rican species; (5) South America clade, consisting primarily of South American endemic species. The latter two clades are sister to each other and together form the most species-rich and morphological diverse clade in

the genus, referred to here as the Central-meridional clade. All the above clades were strongly supported by all analyses except for the Mesoamerica clade that was also resolved by all analyses albeit with weaker support (MP=57.8%; ML=36%; BI=0.63).

CONCATENATED MATRIX VS. SPECIES TREE INFERENCE ANALYSES

Although all phylogenetic methods resolved the same broad topologies and recovered the same major clades (fig. 3), the species tree analysis recovered several relationships more congruent with morphological similarities and/or geographic proximity than those found in the concatenated analyses. The species tree will therefore be used in this paper to discuss phylogenetic relationships, node age estimates, and biogeographic patterns within the genus *Zamia*. The MP (fig. B1) and ML (fig. B2) trees are presented in the supplementary material (available online) and will be referred to primarily in the context of topological and node support congruency with the maximum clade consensus (MCC) species tree topology (fig. 4).

TIME-CALIBRATED SPECIES TREE

Single-locus MCC trees produced by the *Beast analysis generally yielded poorly resolved trees (figs. C3–C12 are available online) compared to the species tree topology (fig. 4). The Effective Sample Size (ESS) scores from the time-calibrated species tree analysis were all over 200, indicating good levels of convergence of the MCMC. The mean stem age of *Zamia* was calculated at 68.28 Ma (95% HPD 51.0–84.5) and the crown age at 9.54 Ma (95% HPD 9.0–10.62) (fig. 5, table 6). The crown node that includes the Mainland *Zamia* clade is the oldest within the genus, with a mean age of 5.97 Ma (95% HPD 3.9–8.2). The Caribbean clade, despite being sister to the rest of the genus, has the youngest crown age 1.19 Ma (95% HPD 0.61–1.86) of the major clades. The Fischeri clade, despite consisting of only three species, is relatively old, with a crown age of 2.59 Ma (95% HPD 0.75–4.62). The Mesoamerica clade, with a crown age of 3.9 Ma (95% HPD 2.54–5.34) is of comparable age to the Central-meridional clade, which is 3.8 Ma (95% HPD 2.49–5.17). Within the Central-meridional clade, the South America clade, with a crown age clade of 2.62 Ma (95% HPD 1.71–3.56), is slightly older than the Isthmus clade, which is 2.35 Ma (95% HPD 1.43–3.32).

DIVERSIFICATION ANALYSES

Net Diversification rates

Net diversification rates were highest when lower extinction rate (ε) values were used in the calculations (table 8). The net diversification rate for the crown of *Zamia* ranged from 0.22 to 0.39. Among the major clades examined, using ε values of 0 and 0.7, the Caribbean clade had the highest diversification rates (1.26, 1.26, 0.88), followed by the South America clade (1.04, 1.03, 0.81) and the Isthmus clade (0.88, 0.65, 0.37). The lowest diversification rates were found in the species-poor Fischeri clade (0.16, 0.10, 0.05), followed by the Mesoamerica clade (0.60, 0.46, 0.27). In the analysis using ε=0.9, the diversification rate was slightly higher for the South America clade than the Caribbean clade. However, ε values of 0.9 are exceptionally high (Magallon & Sanderson, 2001) and may be unrealistic for such a short timeframe, so the results using ε = 0.9 should be interpreted with caution.

Per million year diversification and diversification rate heterogeneity

Most diversification above the crown node of *Zamia* occurred in the Pliocene and Pleistocene epochs (figs. 6, D1), with only the split between the Caribbean and Mainland clade and between the Fischeri clade and the rest of the Mainland clade occurring in the Miocene. The highest per-million year diversification rates were found in the Caribbean clade followed by the South America clade (fig. 6, table 9), the two clades that also had the highest net diversification rates. Among the major *Zamia* clades, the MCCR tests only found diversification rates differing significantly from constant ($P<0.05$) in the South America clade. In this case, the negative γ statistic indicated rapid, early diversification of the South America clade followed by a decrease in diversification rate over time. Similarly, the BAMM analysis did not uncover shifts in diversification rates across the tree. Lastly, the TESS analysis showed no evidence of shifts in extinction or diversification rates, and only marginal evidence of mass extinctions at around 6 and 2.5 mya (fig. D2, available online).

ANCESTRAL CHARACTER STATE RECONSTRUCTION

All four morphological traits examined exhibited homoplasy, and the ancestral states recovered include subterranean stems, smooth petioles, entire leaflet margins, and leaflets without prominent veins

(figs. E1-E4, available online). Arborescence (fig. E1) is absent in the Caribbean clade and rare in the Mesoamerican clade where it is restricted to only three species in the Tuerckheimii clade. The ancestral state for the Central-meridional clade is ambiguous, leading to an ancestral state of arborescence for the Isthmus clade and subterranean stems for the South America clade. Arborescence is most prevalent in the Isthmus clade, with only two species possessing subterranean stems. Within the South America clade, arborescence seems to have evolved separately four times. Regarding petiole armature, unarmed petioles are universal within the Caribbean clade, but extremely rare among mainland species (fig. E2). Both leaflet margin dentation (fig. E3) and leaf venation prominence (fig. E4) are extremely homoplastic characters that appear to have evolved independently multiple times within the genus and may be of limited use as diagnostic characters.

BIOGEOGRAPHIC ANALYSES

Historical biogeography of Zamia

Optimization of ancestral areas on the species tree (fig. 7) required six dispersal and five vicariance events (table 7) with most speciation events occurring within geographical areas. Dispersal events occurred from the Mesoamerica (B) region to the Isthmus (C) and South America (D) regions, from the Isthmus to the South America region, and from South America to the Isthmus region.

The crown node of *Zamia* + *Microcycas* (Node 152, probability of ancestral range at node (*P)*=0.52) is optimized in two areas, one being the Caribbean (A) region, the other being a composite area including the Caribbean (A) and Mesoamerican (B) regions with a dispersal event from the composite area to the Caribbean region. The crown node of *Zamia* (151, *P*=1) is optimized in a composite area consisting of the Caribbean (A) and Mesoamerica (B) regions with a vicariance between these two regions separating the Caribbean clade from the rest of the genus on the American mainland. The Caribbean clade is ancestrally optimized in the Caribbean (A) region with no events hypothesized (85, P=1) and stasis within all inner nodes (78–84, all with *P*=1). The crown node of the clade including all mainland zamias (150, P=0.63) is optimized in the Mesoamerican region (B) with no events hypothesized. The crown node of the Fischeri clade is optimized to the Mesoamerica region (87, *P*=1) with no events hypothesized. The clade including

all *Zamia* except the Caribbean and Fischeri clade (149, *P*=0.21) has a poorly supported optimization among four areas, with the highest support value found for the Mesoamerican region (B, 62.6%). The other three areas were all composite areas that included the Mesoamerica region. Two dispersal events are hypothesized from the Mesoamerica region to the Isthmus and South America regions. The Mesoamerica clade (104, P=1) is optimized in the Mesoamerican region with high probability (*P*=1), no events hypothesized and stasis within all inner nodes (88–103, all with *P*=1). The node that includes *Zamia soconuscensis* as sister to the Central-meridional clade has a poorly supported optimization (148, *P*=0.34) comprised of three composite areas, all of which include the Mesoamerican region and vicariance between Mesoamerica and a composite area consisting of the Isthmus and South America regions.

The split of the Central-meridional clade (147, *P*=0.98) into the Isthmus and South America clades was attributed to a vicariance event occurring 3.8 Ma (95% HPD 2.49–5.17) from a well-supported (*P*=0.98) ancestral composite area including the Central American Isthmus and South America. The Isthmus clade (146, *P*=1) is optimized in that region with no events hypothesized. Stasis is maintained in all inner nodes except the crown node including *Z. obliqua*, *Z. stevensonii*, and *Z. elegantissima*. This node is optimized for the Isthmus (C) region with one dispersal event of *Z. obliqua* into the South America (D) region. The South America clade (131, *P*=0.98) is mostly optimized for that region (D, 97.5%) with no events hypothesized and all but a few inner nodes in stasis. Of note is the clade (110, P=1) that includes *Z. pyrophylla* Calonje, D.W.Stev. & A.Lindstr., *Z. cunaria*, and *Z. ipetiensis* D.W.Stev., which is optimized to a composite area including the Isthmus (C) and South America (D) regions with one vicariance event between these two regions.

Mainland Zamia *distribution gap*

The MaxEnt distribution models for the Isthmus and Mesoamerica clades, and for *Zamia soconuscensis* all had excellent performance, with AUC values of 0.982, 0.967, and 1.0, respectively. The environmental variables contributing most to the models differed for the two clades and *Z. soconuscensis,* with temperature variables being more important for the two clades, and precipitation variables more important for *Z. soconuscensis* (table 10). For the Isthmus clade, mean diurnal temperature range had the greatest percent contribution (66.6%) and permutation importance (41.2%) in the model. For the

Mesoamerica clade, temperature seasonality contributed the most in terms of percent contribution (36.6%) and permutation importance (48.6%). For *Z. soconuscensis*, precipitation in the wettest month (36.3%) and precipitation in the coldest quarter (27.9%) had the greatest percent contribution, whereas precipitation in the coldest quarter (98.4%) was the greatest contributor to permutation importance. The potential distribution maps for both clades present little overlap (fig. 8), with most of the area separating both clades appearing unsuitable for either clade. Suitable habitat for *Z. soconuscensis* extends along the west of the Sierra Madre de Chiapas into Guatemala and El Salvador (fig. 8).

Biogeographic arrangement of genome sizes in Zamia

We recovered a strong phylogenetic signal on the evolution of genome size within the crown group of *Zamia* (Pagel's lambda = 0.9721436), with both genome expansion and contraction occurring within different lineages (fig. F1, available online).

DISCUSSION

UTILITY OF PSBK/I FOR BARCODING AND PHYLOGENETIC ANALYSES

Despite the high rate of amplification success across all cycad genera and a high number of diagnostic sites for DNA barcoding demonstrated for *psbK/I* (Nicolalde-Morejón *et al.*, 2011), its utility for barcoding and phylogenetic inference may be limited by the fact that sizeable inversions are present in several taxa. If the inversions are not detected or dealt with during multiple sequence alignment, they could prove problematic because they would result in multiple false positive single nucleotide polymorphisms. For example, the unusually high number of plant bar coding diagnostic sites reported for *Zamia pseudoparasitica* by (Nicolalde-Morejón *et al.*, 2010) appears to be due to an inversion present in this species that was neither reverse-complemented nor staggered during the alignment process. A comparable situation has also been found in the chloroplast barcoding candidate locus *trnH-psbA* where frequent inversions in multiple angiosperm lineages present a challenge to the gene's utility as a plant barcoding region (Whitlock, Hale & Groff, 2010).

PHYLOGENETIC RELATIONSHIPS AND MAJOR CLADES WITHIN ZAMIA

Below we consider the phylogenetic relationships within *Zamia*, providing brief descriptions of the major clades recovered in our analyses as well as for some of their notable subclades. Informal names are provided for these monophyletic groups for the purposes of discussion clarity (fig. 4).

Monophyly of Zamia

The monophyly of *Zamia* was strongly supported (fig. 4, node A, BI/MP/ML = 1/100/100) and its sister relationship to *Microcycas* was recovered in all analyses. *Zamia* consists of two sister clades: the Caribbean clade, consisting primarily of Caribbean island species, and the Mainland clade, which is largely comprised of mainland American species.

Caribbean clade

The Caribbean clade (fig. 4, node A, BI/MP/ML = 1/100/100) consists of nine species with subterranean-stems, unarmed petioles, and the same 2n=16 chromosome number (Olson & Gorelick, 2011). The group is primarily distributed on several islands in the Greater Antilles and Bahama Archipelago, with a single mainland species restricted mostly to Florida, U.S.A. albeit with a few historical collections in extreme Southeast Georgia (Duncan, 1979). Caribbean zamias are universally accepted as a distinct lineage (e.g. Eckenwalder, 1980a, Sabato, 1990, Stevenson, 1987), yet their current classification (e.g. Stevenson, 1987, González-Géigel, 2003, Osborne *et al.*, 2012) remains controversial. This classification is primarily based on a historic over-reliance on vegetative morphological characters, especially leaflet macromorphology, for species delimitations. Consequently, populations exhibiting similar leaflet morphologies are currently considered conspecific (González-Géigel, 2003) despite sometimes occurring on geographically distant islands that in many cases were never connected by land. This scenario would require multiple occurrences of long-distance dispersal over the Caribbean Sea. An alternative explanation is that some of these morphotypes may have evolved independently multiple times within the clade. Recent genetic studies have begun to address this question and provided further insights into the genetic relationships within this group (Meerow *et al.*, 2012, Meerow *et al.*, in press, Salas-Leiva *et al.*, 2017,

Meerow, 2007). On-going systematics research should help clarify the currently controversial taxonomy of Caribbean zamias in the near future.

Mainland clade

The Mainland clade (fig. 4, node B, BI/MP/ML = 0.95/100/100) is sister to the Caribbean clade and consists chiefly of species occurring on the American mainland, with only a few species also occurring on near-shore continental islands (e.g. *Zamia hamannii* A.S.Taylor, J.L.Haynes & Holzman). It is the most species-rich group in *Zamia* (90% of species) and has a broad geographic range that extends from Mexico to Bolivia. It is comprised of multiple geographically-defined clades discussed below.

Fischeri clade

The Fischeri clade (fig. 4, node C, BI/MP/ML = 100/1/1) is sister to all remaining mainland species and is comprised of three species occurring in northeastern Mexico. *Zamia fischeri* Miq. and *Z. vazquezii* D.W.Stev., Sabato & De Luca are morphologically very similar with the latter segregated from the former based on differences in leaf size, leaflet shape and texture, and differing chromosome counts (Stevenson *et al.*, 1998). Our analyses confirm the two species form a monophyletic group, and we report for the first time this group's sister relationship to *Z. inermis* Vovides, J.D.Rees & Vázq.Torres. All three species occur in relative geographic proximity, have unarmed or sparsely armed petioles, and stems that are subterranean or very short. In addition, *Z. inermis* shares the same chromosome count (2n=16) as *Z. fischeri* (Vovides, 1983, Stevenson *et al.*, 1998) and has a similar reproductive phenology to *Z. vazquezii* (Griffith *et al.*, 2012).

The Mesoamerica clade

This group (fig. 4, node D) of 18 species is restricted to the Mesoamerican dominion (*sensu* Morrone, 2014) which it shares with *Z. soconu*scensis and the three Fischeri clade species. The clade was recovered in all analyses, albeit with low support (below 70%). It is divided into the sister Tuerckheimii and Mexico subclades. The Tuerckheimii subclade (fig. 4, node E), recovered only in the species tree analysis with high support (BI=0.96), is a group comprised of six species endemic to Guatemala, Honduras, and Belize. The Mexico subclade (fig. 4, node B, BI/MP/ML = 1/98.4/98) is comprised of mostly Mexican endemic species

with subterranean stems, armed petioles, and dentate leaflet margins. This group is further divided into sister Purpurea and Furfuracea subclades. Although the Purpurea subclade (fig. 4, node G) was only poorly supported in the species tree analysis (BI=0.64), its constituent species have been recognized as morphologically distinct by several authors (e.g. Schutzman, 1984, Schutzman & Vovides, 1998, Pérez-Farrera *et al.*, 2012), albeit always with the inclusion of *Z. standleyi* Schutzman, which in our results was resolved as deeply embedded within the Tuerckheimii subclade (fig. 4, node E). *Zamia katzeriana* (Regel) E.Rettig, considered by Pérez-Farrera *et al.* (2016) to be a species of hybrid origin, may belong in this group (see Nicolalde-Morejón *et al.*, 2008) but was not included in our sampling. The Furfuracea clade (fig. 4, node H, BI/MP = 0.89/52.4) is comprised of subterranean-stemmed species, most of which are endemic to Mexico with a few species extending into neighboring countries.

Phylogenetic placement of Zamia soconuscensis

The Mexican species *Zamia soconuscensis* was recovered without close relatives in Mesoamerica, but rather as sister to the large Central-meridional clade from the Central American Isthmus region and South America. This inclusive clade was strongly supported in the concatenated analyses, but poorly in the species tree (BI/MP/ML = 0.4/98/89). The Soconusco region in Chiapas where *Z. soconuscensis* occurs is believed to have been a primary Pleistocene floristic refuge (Toledo, 1982), so this species may be the sole survivor of a once larger group of species that were eliminated from surrounding areas during Pleistocene glaciations. Alternatively, the species may have other close relatives in neighboring Guatemala, as on-site observations point to several as yet undescribed species of *Zamia* in the region. The potential discovery of related species and/or the inclusion of additional loci may aid in phylogenetic reconstructions to explain or reject the unusual placement recovered for *Z. soconuscensis* in our analyses.

Central-meridional clade

The Central-meridional clade (fig. 4, node I, BI/MP/ML = 0.98/99.6/95) includes over 60% of all species in the genus, making it by far the most species-rich clade within the mainland zamias. It includes all species occurring from southern Nicaragua south to the southern limit of the genus's range in Brazil and Bolivia. It is separated by a large distribution gap from Mesoamerican *Zamia* populations, which only extend to northern Honduras and El Salvador. This group is comprised of two sister clades: one consisting

mainly of species endemic to the Central American Isthmus (Isthmus clade), and another consisting mainly of South American species (South America clade).

Isthmus clade

The Isthmus clade (fig. 4, node J, BI/MP/ML = 1/99.9/99) consists of fifteen species mostly restricted to the Central American Isthmus region, with only *Zamia obliqua* extending into Colombia in adjacent South America. Most species are arborescent, with only two species (*Z. nana* and *Z. dressleri*) having subterranean stems. Notable groups within the Isthmus clade include the Acuminata subclade (fig. 4, node K, BI=0.92) with four species restricted to the Pacific side of the Cordillera Central mountain range in Costa Rica and Panama, the Obliqua subclade (fig. 4, node L, BI/ML = 0.88/68), comprised of arborescent species with very sparsely armed petioles, and the Skinneri subclade (fig. 4, node M, BI/MP/ML = 0.83/100/100), composed of five arborescent species with large, broad, prominently-veined leaflets and occurring primarily on the Atlantic side of the Cordillera Central from central Panama through southern Nicaragua. *Zamia dressleri*, a species with similar leaflet morphology but with a subterranean stem, has previously been considered part of the *Z. skinneri* Warsz. ex A.Dietr. species complex (Taylor *et al.*, 2008), but this relationship was not supported by our analyses which recovered the species as sister to the Obliqua subclade.

South America clade

The South America clade (fig. 4, node N, BI/MP/ML = 1/100/98) is comprised of twenty-eight species that are endemic to South America, with the exception of *Zamia manicata* Linden ex Regel which extends into Panama, and sister species *Z. cunaria* and *Z. ipetiensis* (Cunaria clade, fig. 4, node R, BI= 0.89), which are Panamanian endemics. The phylogenetic relationship of the latter two species to other species occurring in South America remains unclear, as the species tree and concatenated analyses resulted in conflicting and poorly supported topologies. However, these species share extensive morphological similarities with *Z. pyrophylla* from the Colombian Chocó, including subterranean stems holding only one to two leaves, tomentum present on strobili axes, and microsporangia on the adaxial side of the microsporophylls (Calonje *et al.*, 2010).

The Manicata subclade (fig. 4, node O, BI/ML/BI = 1/99.3/98) is a group of four species occurring in primarily in northern Colombia that have a remarkably variable leaflet morphology including the distinctly channeled petiolule with a gland-like collar of *Zamia manicata*, the membranaceous and prominently veined leaflets of *Z. disodon* D.W.Stev. & Sabato, and the unusual midrib found in *Z. restrepoi* (D.W.Stev.) A.Lindstr. The midrib found in *Z. restrepoi* leaflets is unique in the genus and so distinctive that the species was initially described within its own genus (*Chigua* D.W.Stev.) until it was eventually subsumed into *Zamia* (Lindstrom, 2009). The strikingly diverse leaflet morphology found in this clade has historically obscured the close phylogenetic relationships of its constituent species, but they do share other characters such as acaulescent stems, toothed leaflets, relatively small seeds, and microsporophylls with a very short fertile section of lamina; these characters support the close relationship recovered in our analyses.

The Pacific subclade (fig. 4, node P, BI/ML/BI = 0.78/70/61) is comprised of five species occurring from the western foothills of the Andes to the Pacific coast. The species are variable in terms of leaflet vein prominence, stem habit, and leaflet margin dentation (figs. E1–E4, available online) but are all extensively armed with prickles on the petioles, making it the most aggressively armed clade in the genus. Panamanian highland species *Z. lindleyi* Warsz. Ex A.Dietr. was previously considered conspecific with *Z. chigua* Seem. (Stevenson, 1993) from the lowlands of the Colombian Chocó biogeographic region primarily due to having fern-like leaves with numerous narrow leaflets. However, several morphological and ecological differences in addition to the large geographic disjunction have since brought recognition of *Z. lindleyi* as a distinct and separate species from *Z. chigua* (Calonje *et al.*, 2012b), a recognition supported by our analyses, which place the former within the Isthmus clade, and the latter in the South America clade. Norstog (1980) believed that Pacific subclade species *Z. roezlii* Linden (as *Z. chigua*; Norstog, 1986) was the most primitive species in the genus due to its stable rainforest habitat allowing it to retain what he considered primitive characters such as an arborescent habit, large spermatozoids and a large asymmetrical karyotype. However, three of the four traits we examined for this species, including the arborescent habit, were not ancestral in our character state reconstructions (figs. E1–E4). Additionally, karyotype evolution in *Zamia* appears to be moving towards increased asymmetry with higher numbers of smaller chromosomes

(Olson & Gorelick, 2011, Vovides & Olivares, 1996), meaning the large asymmetric karyotype of *Z. roezlii* is currently considered more derived than ancestral.

The Wallisii subclade (fig. 4, node Q, BI/MP/ML = 0.78/93/93) is comprised of *Zamia wallisii* Braun and *Z. oligodonta* E. Calderón & D.W.Stev., two subterranean-stemmed, montane species with prominently-veined, extremely broad, coriaceaous leaflets and ovoid female cones with few, relatively large seeds, both occurring on the western flank of Colombia's Western Cordillera. *Zamia montana* Braun, an arborescent species with prominently veined leaflets also has prominently veined leaflets and occurs in the same geographical region. It was not sampled in this study but likely belongs to this subclade as well. In fact, *Z. montana* and *Z. oligodonta* were previously considered conspecific due to vegetative similarities (Lindstrom, 2009), but recent field studies with the two species (Calonje *et al.*, 2015) has clarified that the two species are significantly morphologically distinct (Calonje *et al.*, 2015).

The Tolimensis subclade (fig. 4, node S, BI = 0.67) consists of sister species, *Zamia tolimensis* Calonje, H.E.Esquivel & D.W.Stev. and *Z. huilensis* Calonje, H.E.Esquivel & D.W.Stev., both endemic to Colombia and occurring in mountain ranges surrounding the Magdalena river valley. The former occurs in the Central Cordillera, the latter in the Eastern cordillera and the Colombian Massif. Both are arborescent species sharing considerable similarities in reproductive structures (Calonje *et al.*, 2012a).

The Eastern subclade (fig. 4, node T, BI/MP/ML = 0.81/64/71) is comprised of species occurring in Eastern Colombia and Western Venezuela. All species occur to the east of the Andes mountain range with the exception of *Z. encephalartoides* D.W.Stev., which occurs on the western side of the Cordillera Oriental mountain range. The taxon treated here as *Z. lecontei* Ducke is from Amazonian Venezuela, following the species circumscription by Stevenson (2004) . However, the type locality for this species is a great distance away in Brazil, and additional taxonomic work is required to determine if Brazilian, Venezuelan, and Colombian populations of this species are truly conspecific.

The Amazonian subclade (fig. 4, node U, BI/MP/ML =0.64/70/66), recovered in all analyses (albeit with poor support), consists of seven species endemic to the Amazon basin. *Zamia poeppigiana* Mart. & Eichler, a large arborescent Amazonian species, shares many morphological similarities with *Z. lindenii*

Regel ex André, which occurs on the Pacific side of the Andes. The two were previously considered conspecific (Stevenson, 2001) until additional taxonomic work helped clarify morphological differences between them (Lindstrom, 2009, Calonje *et al.*, 2011). Their placement in the separate Pacific and Amazonian subclades in our analyses justify their recognition as separate species.

ANCESTRAL CHARACTER STATES AND HOMOPLASY OF MACROMORPHOLOGICAL VEGETATIVE CHARACTERS

Homoplasy was prevalent within all the macromorphological characters examined (i.e. arborescence, petiole armature, leaflet margin dentation, and leaflet vein prominence), and no single character or combination thereof appeared to be diagnostic for the major clades within *Zamia* (Figs. E1-E4, available online). This agrees with previous findings by Caputo *et al.* (2004), who found all characters from an extensive morphological and micromolecular dataset to be homoplastic, as well as inferences from population based studies in the Caribbean clade (Meerow *et al.*, in press, Salas-Leiva *et al.*, 2017). The rampant homoplasy pervasive in the genus has resulted in several species that share morphological resemblance but are not closely related. This is perhaps best exemplified by species with prominently veined leaflets, a character unique to the genus and so distinctive that it was the primary reason for the segregation of the genus *Aulacophyllum* Regel from *Zamia* (Regel, 1876). This character state, found in approximately one quarter of all *Zamia* species, has evolved multiple times throughout the genus and is not diagnostic of any clade (fig. E4, available online). Similarly, several large arborescent South American species with prominently toothed leaflet margins (*Z. tolimensis*, *Z. lindenii*, and *Z. poeppigiana*), which were previously considered to belong to the same species complex (Calonje *et al.*, 2011), are actually members of three separate lineages within the South America clade. The pervasiveness of homoplasy within macromorphological vegetative characters typically considered diagnostic complicates their diagnostic utility, and future systematic research will need to consider a wide array of characters, including morphological, anatomical, reproductive, and genetic characters.

DIVERGENCE AGE ESTIMATION

The divergence between *Zamia* and *Microcycas* was estimated to be 68.28 Ma (95% HPD 50.99–84.5), an estimate older than those presented by Condamine et al. (57 Ma), Salas et al. (36.5 Ma), and Nagalingum et al. (31.1 Ma). Our estimate suggests that the divergence between the two genera happened near the Cretaceous-Paleogene boundary which was marked by the catastrophic Chicxulub asteroid impact that may have wiped out up to 60–70% of plant species in tropical North America (Nichols & Johnson, 2009). The impact likely had a devastating effect on contemporary cycad flora in the Americas, but the aftermath likely presented a speciation opportunity for surviving cycad genera, and in fact the Cycadales as a whole experienced major diversification during this time (Salas-Leiva *et al.*, 2013). The estimated stem age for *Microcycas* is older than the present Antillean islands, which were formed less than 40 Ma (i.e., late Eocene) (Iturralde-Vinent, 2006). This suggests the ancestral range of *Microcycas* may have originally included mainland populations that eventually became extinct. However, such interpretation should be taken cautiously as the fossil record for *Zamia* is scarce and is, as yet, inexistent for *Microcycas.* As the latter is a monotypic genus, there are no diversification events to infer that could clarify its biogeographic history.

The mean crown age of *Zamia* was estimated at 9.54 Ma (95% HPD 9.0–10.6 Ma). This age is younger than reported by Condamine *et al.* (2015) (14.6 Ma, using new fossil dataset and birth/death prior), slightly younger than reported by Nagalingum et al. (11.25 Ma, Bayesian inference using nuclear markers), and slightly older than reported by Salas et al. (8.2 Ma). Our crown age estimate for *Zamia* marks the divergence of the Caribbean clade from the Mainland clade, and is concordant with the time-since-divergence estimate (Sarich, 1977) of 10.8 Ma between *Z. pumila* L. (*sensu* Eckenwalder, 1980a) and *Z. splendens* reported by Walters and Decker Walters (1991) based on allozyme differentiation.

BIOGEOGRAPHY OF ZAMIA

Geographic delimitation of clades

Three separate major areas of endemism are readily identifiable in the geographic distribution of the genus *Zamia*: the Caribbean Islands and Southeast U.S.A, Mesoamerica, and Central+ South America (fig.

1). The *Zamia* species occurring in the Caribbean Islands and Southeast U.S.A. are all monophyletic (Caribbean clade), as are all species from Central and South America (Central-meridional clade). However, in the Mesoamerica region we found two separate clades with overlapping geographic distributions (Fischeri and Mesoamerica clades) as well as the unusual situation with *Z. soconuscensis,* which in our analyses appears more closely related to the Central-meridional clade than to any Mesoamerican species. With the notable exception of *Z. soconuscensis*, we found a strong pattern of phylogenetic congruence with geographic distribution in *Zamia*, with most of the major clades recovered in our analyses exclusively occurring in particular geographic regions with little or no range overlap. Indeed, our results bolster similar findings by Caputo *et al.* (2004) by demonstrating that this pattern is widespread throughout the entire genus.

The strong congruence between geographic distribution and phylogenetic relationships recovered in our analysis is a pattern consistent with the limited dispersal and pollen transfer abilities found in cycads (Tang, 1993, Yang & Meerow, 1996), which most likely limit the ability of these clades to expand beyond geographically contiguous and ecologically suitable conditions following isolation via vicariance or dispersal events. Although rare, long-range dispersal events have occurred, as evidenced by the distribution of Caribbean clade species on landmasses that have never had land connections between them, and these events have played a role in shaping the evolutionary history of the genus (Salas-Leiva *et al.*, 2017). Eckenwalder (1980b) posited that overwater dispersal of Caribbean zamias could occur via beached rafts of drift material carried by strong currents or during hurricanes. Additionally, several authors have reported birds as short-range dispersal agents of *Zamia* (Gómez, 1993, Eckenwalder, 1980b, Tang, 1993, Chemnick, 2007). Long-range dispersal could also be possible via larger birds that are able to swallow whole seeds and disperse them via defecation, such as the great Curassow (*Crax rubra* L.) a known dispersal agent of *Ceratozamia whitelockiana* Chemnick & T.J.Greg. (Chemnick, 2007). The distribution range of *Zamia*, by far the broadest in the New World cycads and including several landmasses that were never connected by land, suggests that its dispersal abilities have played an important role in its evolutionary history.

Historical biogeography of Zamia

Inferring the historical biogeography of *Zamia* with any certainty is burdened by an extremely poor fossil record consisting of a single fossil that can be unequivocally assigned to *Zamia* based on morphological and epidermal characters (Erdei *et al.*, 2018). The fossil, estimated to be of late Eocene to earliest Oligocene age (ca. 33–35 Ma), was found in marine sediments in Central Panama. Little is known about its terrestrial origin, but since parts of the Panama Arc were emergent since at least 30 Ma (O'Dea *et al.*, 2016), it is possible that leaflets may have been carried by currents from nearby islands or more distant land. Interestingly, the cuticular micromorphology of the Panamanian fossil more closely resembles that of extant Caribbean clade species than that of Central-meridional clade species (Erdei *et al.*, 2018). This may imply that the elongated adaxial cuticular cell shape, and stomatal band width and ratio found in Caribbean species are ancestral character states. Two additional Paleogene (*i.e.*, Oligocene 30 Ma) fossils from Collazo Shale (Puerto Rico) were originally described as separate species of *Zamia* (Hollick, 1928) but are apparently conspecific (Erdei *et al.*, 2018). Despite their *Zamia*-like macromorphology, their confirmation as *Zamia* is not possible due to poorly preserved cuticles (Erdei *et al.*, 2018). If assignable to *Zamia*, the fossils would indicate the earlier presence of the genus in the Caribbean, albeit likely of an extinct lineage, as Puerto Rico was below sea level between the Late Oligocene and the Early Pliocene (van Gestel *et al.*, 1998). Hence, the current presence of *Zamia* on the island may be younger than ca. 5 Ma (Meerow *et al.*, 2012), which is in agreement with our estimations that suggest that the diversification of extant Caribbean species occurred in less than 2 Ma, and within the last 340 thousand years for Puerto Rican and Dominican species (fig. 5). Despite having an old stem age (9.54 Ma, 95% HPD 9.0–10.62) dating back to the origin of extant zamias (fig. 5), the Caribbean clade has a very young crown age (1.19 Ma, 95% HPD 0.61–1.86), perhaps the result of high turnover rates in the clade resulting from glacio-eustatic sea level fluctuations, which can greatly influence island biotas (Weigelt *et al.*, 2016). In fact, the effects of historic sea level fluctuations are readily observable in the genetic structure of extant species. Sea level fluctuations are thought to have reinforced isolation in *Zamia erosa* O.F.Cook & G.N.Collins limestone hill populations in northern Puerto Rico by eliminating intervening populations in the lowlands (Meerow *et al.*, 2012). Similarly, populations of *Zamia* on five islands in the Bahamian archipelago cluster genetically into two

groups (Salas-Leiva *et al.*, 2017) that correspond to two separate historic landmasses (Little Bahamas Bank and Great Bahamas Bank) which encompassed these islands during the Last Glacial Maximum (33.0 and 26.5 Ka) when sea levels were as much as 120 to 135 m below current levels (Clark & Mix, 2002).

Based on the limited fossil record for *Zamia*, Erdei *et al.* (2018) suggested an Eocene or earliest Oligocene origin of the genus in the Central American – Caribbean region with subsequent dispersal south and northwards. Our molecular biogeographic and divergence time estimation analyses instead suggest an older Late Cretaceous to early Eocene origin for the stem group of *Zamia*, and a Caribbean or Mesoamerican origin for both the stem and crown nodes of *Zamia* with subsequent dispersal of the genus south into the Central American Isthmus region and South America. Despite the discovery in the Central American Isthmus region of the only anatomically-verified *Zamia* fossil (albeit of unknown terrestrial provenance), a Caribbean or Mesoamerican origin for the genus appears more convincing due to the presence of the sister genus *Microcycas* as well as the two earliest diverging *Zamia* lineages (Caribbean and Fischeri clades) at the northernmost range of the genus's geographic range. Our biogeographic analysis suggests a northern origin for the crown group of *Zamia* with vicariance caused by a southward distribution shift in the Late Miocene, a sub-epoch characterized by increasing drought, enhanced seasonality, and a restructuring of terrestrial biological communities (Herbert *et al.*, 2016). Discoveries of additional fossils of *Zamia* or it sister genus *Microcycas* may help better recreate the still poorly understood early biogeographic history of the genus.

There is currently considerable disagreement regarding dating of the emergence of the Panama land bridge, with several authors advocating the traditional view of ca. 2.8 Ma (O'Dea *et al.*, 2016) and others suggesting the isthmus has been in place for over 10 Ma (Jaramillo *et al.*, 2017, Montes *et al.*, 2012). Either way, the estimated divergence of the Central-meridional clade into the Isthmus and South America clades (3.8 Ma, 95% HPD 2.49–5.17)) suggests that the two groups diverged when the Panama land bridge was either nearly completed (O'Dea *et al.*, 2016) or had been in place for millions of years.

Diversification of the crown nodes for both clades began nearly simultaneously about a million years later, with the South America clade's mean crown age estimated at 2.62 Ma (95% HPD 1.71–3.56) and the

Isthmus clade's at 2.35 Ma (95% HPD 1.43–3.32). This period coincides with the establishment of repeated Pleistocene glaciations beginning at 2.6 Ma which converted much of Central America into recurrently dry environments (Bacon *et al.*, 2016) leading to the fragmentation of wet forest habitat favored by the Central-meridional clade and likely driving diversification of both the Isthmus and South America clades.

The South America clade is considerably more species-rich and has a broader geographic range than the Isthmus clade, perhaps in part because the dispersal potential of the latter has been limited by the geographic bottleneck of the Isthmus's narrow landmass. Despite this bottleneck the Isthmus clade has been remarkably successful. It has the third highest net diversification rate after the Caribbean and South America clades (table 8), and the Isthmus region hosts the greatest concentration of species per unit area in the genus, with the main taxonomic diversity hotspots occurring in central Panama, and along the Costa Rican and Panamanian border (fig. 9).

The Isthmus and South America clades appear to have diversified and dispersed mostly within their geographic regions despite the fact that this diversification occurred at when the Panama land bridge was already present or near-completed. The genetic isolation of these clades mirrors similar findings reported for birds (Smith & Klicka, 2010) and mammals (Webb, 1976) in which biotic interchange facilitated by uplift of the isthmus declined during the Pleistocene perhaps driven by the unavailability of ecological space. Reduced migration by birds and mammals, both known dispersal agents of *Zamia* (Tang, 1989) may have affected trans-isthmian dispersal of the genus, and the ecological niches available to *Zamia* may have also filled rapidly in both landmasses, therefore impeding biotic exchange. Lastly, the co-evolution of *Zamia* species with insect pollinators also may have played a role in keeping the clades separate. In fact, whereas erotylid beetle pollinators of the genus *Pharaxonotha* Reitter are found throughout most of the distribution range of *Zamia*, weevil pollinators of the subtribe *Allocorynina* Sharp are absent from South America clade species (O'Brien & Tang, 2015).

Biogeographic arrangement of genome sizes in Zamia

We recovered a strong phylogenetic signal for the evolution of genome size within the crown group of *Zamia* (Pagel's lambda = 0.9721436), indicating a strong statistical dependence between DNA content and

our species tree topology. Genome size appears to have evolved stochastically within the genus, with both increases and decreases in DNA content occurring in different clades. Zonneveld and Lindström (2016) noted patterns between genome size and geographic distribution, identifying three major biogeographic groups matching the Mesoamerican, Caribbean, and South + Central American areas of endemism (fig. 1) for the genus. However, the latter two groups are not diagnosable based on genome size alone, as they include overlapping values, a situation that Zonneveld and Lindström (2016) attributed to ancient hybridizations between the two groups. Our results instead suggest that these similarities are due to both groups having genome sizes that have remained stable since the diversification of the crown node of the genus (fig. F1, available online). Similarly, the Mesoamerican group, while diagnosable by its low genome size, in our phylogenetic analyses represents a paraphyletic assembly that includes members of the separate Mesoamerica and Fischeri clades as well as *Zamia soconuscensis*.

The Honduran-Nicaraguan distribution gap

There is a large gap in the distribution of the genus found from Southern Honduras and El Salvador to Southern Nicaragua (fig. 8a, 8b), with *Zamia soconuscensis*, and the Fischeri and Mesoamerica clades occurring exclusively to the North of this divide, and the Central-meridional clade occurring to the South of the divide. The majority of the diversity and distribution range of *Zamia* in Mesoamerica occurs in the Atlantic side of the Sierra Madre mountain system in Mexico and the Guatemalan, Honduran, and Nicaraguan highlands, with only *Zamia paucijuga* Wieland, *Z. herrerae* S.Calderón & Standl., and *Z. soconuscensis* Schutzman, Vovides & Dehgan occurring on the Pacific side. These mountains form a formidable geographic and climatological barrier that appears to have restricted the distribution of Mesoamerican zamias south into the rest of Central America. The Isthmus clade reaches its northernmost distribution in southern Nicaragua with only the single species *Zamia neurophyllidia* D.W.Stev. occurring here. The center of diversity of the Skinneri clade, to which *Z. neurophyllidia* belongs, occurs in western Panama, which is likely where it originated. With a mean age estimate of 0.77 Ma (fig. 5), the clade is rather young and possibly not enough time has elapsed for it to disperse northward into additional suitable habitat found along the Atlantic Nicaraguan coast (fig. 8d). Even accounting for this additional suitable habitat, the potential ranges for both the Mesoamerica and Isthmus clades present little overlap, indicating

both clades have adapted to different climatological conditions. In both cases the environmental variables contributing the most to their distribution models appear to be temperature-based, with temperature seasonality (difference between annual maximum and minimum temperatures) being the most important variable for the Mesoamerica clade, and mean diurnal temperature range (mean difference between daily maximum and minimum temperatures) being the most important variable for the Isthmus clade (Table 10).

DIVERSIFICATION

Most diversification above the crown node of *Zamia* occurred in the Pliocene and Pleistocene Epochs, with only the split between the Caribbean and Mainland clade and between the Fischeri clade and the rest of the Mainland clade occurring in the Miocene (figs. 5, 6). Climatic oscillations during the Pleistocene appear to have had a great effect on the diversification of modern *Zamia* species, as the vast majority appear to be of Pleistocene origin.

The net diversification rates (sensu Magallon & Sanderson, 2001) recovered for the crown group of *Zamia* are among the highest reported for gymnosperm genera by Crisp and Cook (2011), yet comparable to rates obtained using clade age estimates from Nagalingum *et al.* (2011) and Condamine *et al.* (2015) (table 8). The diversification rates were highest in the Caribbean clade, followed by the South America and Isthmus clades, the two major clades constituting the Central-meridional clade (table 8).

The high diversification rate recovered in the Caribbean clade may be partly due to the very rapid generation times found in this clade, as its constituent species are able to reach reproductive age as early as two years from seed (Salas-Leiva *et al.*, 2017), therefore allowing for a faster rate of accumulation of neutral mutations compared to most mainland species. Additionally, the diversification pattern of Caribbean zamias has likely been affected by major glacio-eustatic sea level fluctuations that affected the region's biota during the Quarternary (Iturralde-Vinent, 2006) and are hypothesized to have had a strong effect on the diversification of Caribbean *Zamia* on the Bahamian archipelago (Salas-Leiva *et al.*, 2017). The effect on Caribbean biota caused by Quaternary glaciations and accompanying sea level changes can be attributed to many factors, including changes in climate associated with glaciations as well as fluctuations in habitat size and inter-island distances affecting gene flow.

The South America clade was the only group in which we found a signal for significant early rapid speciation (fig. 6). This could be indicative of an adaptive radiation triggered by the ecological opportunity presented by vacant niches becoming available following the invasion of *Zamia* into the massive landmass of South America. The instant availability of multiple and varied niches throughout a massive land expanse could have resulted in the Central-meridional clade, as well as its constituent South America and Isthmus clades having the highest diversification rates among mainland zamias, despite diversifying more or less contemporaneously with the Mesoamerica clade (fig. 5). Alternatively, the Mesoamerica clade's slower diversification rate may be a consequence of higher extinction rates. In truth, stochastic processes cannot be ruled out as a reason for the disparity in species richness between these clades, as the BAMM analysis recovered homogeneous diversification and extinction rates across the entire tree.

Nagalingum *et al.* (2011) found patterns of early diversification followed by subsequent decrease in diversification rates for all cycad genera, a pattern typical of an adaptive radiation (Rabosky & Hurlbert, 2015) which they attributed to newly available seasonal environments as the ecological opportunity that triggered a global radiation following the mid-Miocene transition. With the exception of the South America clade, in which an early diversification burst occurred (fig. 6), we found rates of diversification within the genus to be constant over time, following a 'museum model' of evolution (Wallace, 1878, Fischer, 1960) with gradual species accumulation coupled with low extinction rates. This model is supported by the BAMM analysis in which no significant diversification rate shifts were detected within *Zamia* or any of its constituent clades by the MCCR test, which indicated a constant diversification rate for the genus and most major clades (fig. 6, table 9). In addition, no significant mass extinction signals were detected in the genus using the Comet model (fig. D2, available online). The stable diversification history apparent in *Zamia* differs from the findings of Yessoufou *et al.* (2014) in the genus *Encephalartos* Lehm., in which a mass extinction event has been postulated to explain a pattern of rate heterogeneity through time. This difference could indicate the ability of *Zamia* species of to track their favorite habitat during periods of cooling-drying, or it could signify different patterns of cooling and aridification between Africa and the Neotropics (Couvreur, 2015).

ACKNOWLEDGMENTS

Funding for field and laboratory research was provided by the National Science Foundation (DEB 1050340), the National Geographic Society (8965-11), the Mohamed Bin Zayed Species Conservation Fund (project number 0925331), the Association for Zoological Horticulture, the Institute of Museum and Library Services grants (MA-05-12-0336-12: Mission Based Collections Planning and MA-30-14-0123-14: Mission Based Collections Stewardship), SOS—Save Our Species (Grant 2012A-035), Tinker Foundation Field Research Grant, Montgomery Botanical Center, the United States Department of Agriculture Agricultural Research Service, Fairchild Tropical Botanic Garden, Charles P & Dorothy Sacher, Dr. Lin Lougheed, and the Christiane Tyson Research Fellowship granted to D.S.-L. We are grateful to the many botanic gardens, herbaria, and horticulturists that generously provided samples for this study (See Appendix A). Kyoko Nakamura provided laboratory training and support to M.C., Fabien Condamine provided helpful guidance regarding age divergence estimation, Miguel Angel Pérez Farrera provided relevant information about Mexican zamias. Finally, Claudia Patricia Gutierrez provided unconditional family support to M.C. during his dissertation work.

LIST OF REFERENCES

Acuña-Castillo, R, Marín-Méndez, W. 2013. Comparative anatomy of leaflets of *Zamia acuminata* and *Z. pseudomonticola* (Zamiaceae) in Costa Rica. *Revista de Biologia Tropical* **61:** 539-546.

Aiello-Lammens, ME, Boria, RA, Radosavljevic, A, Vilela, B, Anderson, RP. 2015. spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* **38:** 541-545.

Bacon, CD, Molnar, P, Antonelli, A, Crawford, AJ, Montes, C, Vallejo-Pareja, MC. 2016. Quaternary glaciation and the Great American Biotic Interchange. *Geology* **44:** 375-378.

Bogler, DJ, Francisco-Ortega, J. 2004. Molecular systematic studies in cycads: Evidence from trnL intron and ITS2 rDNA sequences. *Botanical Review* **70:** 260-273.

Bouckaert, R, Heled, J, Kühnert, D, Vaughan, T, Wu, C-H, Xie, D, Suchard, MA, Rambaut, A, Drummond, AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **10:** e1003537.

Bouckaert, RR, Drummond, AJ. 2017. bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evolutionary Biology* **17:** 42.

Calonje, M, Esquivel, HE, Morales, G, Mora-Lizcano, YA, Stevenson, D. 2012a. A new arborescent species of *Zamia* (Cycadales, Zamiaceae) from the department of Huila, Eastern Cordillera of Colombia. *Caldasia* **34:** 8.

Calonje, M, Esquivel, HE, Stevenson, DW, Calonje, C, Pava, D. 2011. A new arborescent species of *Zamia* from the Central Cordillera of Tolima, Colombia (Cycadales, Zamiaceae), with comments on the *Z. poeppigiana* species complex. *Brittonia* **63:** 442-451.

Calonje, M, Morales, G, López-Gallego, C, Roldán, F. 2015. A taxonomic revision of *Zamia montana* and *Zamia oligodonta*, with notes on their conservation status *Phytotaxa* **192:** 279-289.

Calonje, M, Stevenson, DW, Calonje, C, Ramos, YA, Lindstrom, A. 2010. A new species of *Zamia* from Choco, Colombia (Cycadales, Zamiaceae). *Brittonia* **62:** 80-85.

Calonje, M, Stevenson, DW, Osborne, R. 2019. The World List of Cycads, online edition (http://cycadlist.org) Accessed (Accessed 21 Sep 2019).

Calonje, M, Taylor-Blake, AS, Stevenson, D, Holzman, G, Ramos, YA. 2012b. *Zamia lindleyi*: A misunderstood species from the highlands of western Panama. *Mem. New York Bot. Gard.* **106:** 419-437.

Caputo, P, Cozzolino, S, Gaudio, L, Moretti, A, Stevenson, DW. 1996. Karyology and phylogeny of some Mesoamerican species of *Zamia* (Zamiaceae). *American Journal of Botany* **83:** 1513-1520.

Caputo, P, Cozzolino, S, Luca, PD, Moretti, A, Stevenson, DW. 2004. Molecular phylogeny of *Zamia* (Zamiaceae). In: Walters, TW, Osborne, R eds. *Cycad Classification: Concepts and Recommendations.* Wallingford, United Kingdom: CABI Publishing.

Chaw, SM, Walters, TW, Chang, CC, Hu, SH, Chen, SH. 2005. A phylogeny of cycads (Cycadales) inferred from chloroplast matK gene, trnK intron, and nuclear rDNA ITS region. *Molecular Phylogenetics and Evolution* **37:** 214-234.

Chemnick, J. 2007. Seed dispersal agents of two Mexican cycads. *Mem. New York Bot. Gard.* **97:** 104-119.

Clark, PU, Mix, AC. 2002. Ice sheets and sea level of the Last Glacial Maximum. *Quaternary Science Reviews* **21:** 1-7.

Clugston, JAR, Griffith, MP, Kenicer, GJ, Husby, CE, Calonje, MA, Stevenson, DW, Little, DP. 2016. *Zamia* (Zamiaceae) phenology in a phylogenetic context: does in situ reproductive timing correlate with ancestry? *Edinburgh Journal of Botany* **73:** 345-370.

Condamine, FL, Nagalingum, NS, Marshall, CR, Morlon, H. 2015. Origin and diversification of living cycads: a cautionary tale on the impact of the branching process prior in Bayesian molecular dating. *BMC Evolutionary Biology* **15:** 65.

Couvreur, TLP. 2015. Odd man out: why are there fewer plant species in African rain forests? *Plant Systematics and Evolution* **301:** 1299-1313.

Crisp, MD, Cook, LG. 2011. Cenozoic extinctions account for the low diversity of extant gymnosperms compared with angiosperms. *New Phytologist* **192:** 997-1009.

Duncan, WH. 1979. *Zamia* (Cycadaceae) new for Georgia. *Sida* **8:** 115-116.

Eckenwalder, J. 1980a. Taxonomy of the West Indian cycads. *Journal of the Arnold Arboretum* **61:** 701-722.

Eckenwalder, JE. 1980b. Dispersal of the West Indian cycad, *Zamia pumila* L. *Biotropica* **12:** 79-80.

Erdei, B, Calonje, M, Hendy, A, Espinoza, N. 2018. A review of the Cenozoic fossil record of the genus Zamia L. (Zamiaceae, Cycadales) with recognition of a new species from the late Eocene of Panama - evolution and biogeographic inferences. *Bulletin of Geosciences* **93:** 185-204.

Escalante, T, Morrone, JJ, Rodríguez-Tapia, G. 2013. Biogeographic regions of North American mammals based on endemism. *Biological Journal of the Linnean Society* **110:** 485-499.

Farris, JS, Albert, VA, Källersjö, M, Lipscomb, D, Kluge, AG. 1996. Parsimony jackknifing outperforms neighbor-joining. *Cladistics* **12:** 99-124.

Fischer, AG. 1960. Latitudinal variations in organic diversity. *Evolution* **14:** 64-81.

Foote, M. 2000. Origination and extinction components of taxonomic diversity: general problems. *Paleobiology* **26:** 74-102.

GBIF.org. 2017. GBIF Home Page (http://gbif.org) Accessed 29th August 2017.

Gómez, LD. 1993. Birds as short range seed dispersers of *Zamia fairchildiana* in SW Costa Rica. *Revista Biologia Tropical* **41:** 905-906.

González-Géigel, L. 2003. Zamiaceae. *Flora de la República de Cuba.* Liechtenstein: A. R. Gantner Verlag KG.

Griffith, M, Calonje, M, Stevenson, D, Husby, C, P. Little, D. 2012. Time, place, and relationships: cycad phenology in a phylogenetic and biogeographic context. *Mem. New York Bot. Gard.* **106:** 59-81.

Harmon, LJ, Weir, JT, Brock, CD, Glor, RE, Challenger, W. 2008. GEIGER: investigating evolutionary radiations. *Bioinformatics* **24:** 129-131.

Heled, J, Drummond, AJ. 2010. Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* **27:** 570-580.

Herbert, TD, Lawrence, KT, Tzanova, A, Peterson, LC, Caballero-Gill, R, Kelly, CS. 2016. Late Miocene global cooling and the rise of modern ecosystems. *Nature Geosci* **9:** 843-847.

Hill, KD. 1995. Infrageneric relationships, phylogeny and biogeography of the genus *Cycas* (Cycadaceae). In: Vorster, P ed *CYCAD 93, The 3rd International Conference on Cycad Biology, Proceedings.* Stellenbosch: Cycad Society of South Africa.

Hill, KD. 1996. A taxonomic revision of the genus *Cycas* (Cycadaceae) in Australia. *Telopea* **7:** 1-63.

Hodell, DA, Curtis, JH, Sierro, FJ, Raymo, ME. 2001. Correlation of Late Miocene to Early Pliocene sequences between the Mediterranean and North Atlantic. *Paleoceanography* **16:** 164-178.

Höhna, S, May, MR, Moore, BR. 2016. TESS: an R package for efficiently simulating phylogenetic trees and performing Bayesian inference of lineage diversification rates. *Bioinformatics* **32:** 789-791.

Hollick, A. 1928. Paleobotany of Porto Rico. *Scientific Survey of Porto Rico and the Virgin Islands.*

Iturralde-Vinent, MA. 2006. Meso-Cenozoic Caribbean paleogeography: Implications for the historical biogeography of the region. *International Geology Review* **48:** 791-827.

Jaramillo, C, Montes, C, Cardona, A, Silvestro, D, Antonelli, A, Bacon, CD. 2017. Comment (1) on "Formation of the Isthmus of Panama" by O'Dea et al. *Science advances* **3**.

Johnson, L. 1959. The families of cycads and the Zamiaceae of Australia. *Proceedings of the Linnean Society of New South Wales* **84:** 64-117.

Karger, DN, Conrad, O, Böhner, J, Kawohl, T, Kreft, H, Soria-Auza, RW, Zimmermann, N, Peter Linder, H, Kessler, M 2016. *Climatologies at high resolution for the Earth land surface areas*.

Katoh, K, Standley, DM. 2013. MAFFT Multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30:** 772-780.

Lahaye, R, van der Bank, M, Bogarin, D, Warner, J, Pupulin, F, Gigot, G, Maurin, O, Duthoit, S, Barraclough, TG, Savolainen, V. 2008. DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences* **105:** 2923-2928.

Lindstrom, AJ. 2009. Typification of some species names in *Zamia* L. (Zamiaceae), with an assessment of the status of *Chigua* D. Stev. *Taxon* **58:** 265-270.

Maddison, W, Maddison, D. 2017. Mesquite: a modular system for evolutionary analysis.

Magallon, S, Sanderson, MJ. 2001. Absolute diversification rates in angiosperm clades. *Evolution* **55:** 1762-1780.

May, MR, Höhna, S, Moore, BR. 2016. A Bayesian approach for detecting the impact of mass-extinction events on molecular phylogenies when rates of lineage diversification may vary. *Methods in Ecology and Evolution* **7:** 947-959.

McNeill, J, Barrie, FR, Buck, WR, Demoulin, V, Greuter, W, Hawksworth, DL, Herendeen, PS, Knapp, S, Marhold, K, Prado, J, Reine, WFPHV, Smith, GF, Wiersema, JH, Turland, NJ. 2012. International Code of Nomenclature for algae, fungi and plants (Melbourne Code) adopted by the Eighteenth International Botanical Congress Melbourne, Australia, July 2011. *Regnum Vegetabile.* Bratislava: Koeltz Scientific Books.

Meerow, AW, Francisco-Ortega, J, Calonje, M, Griffith, MP, Ayala-Silva, T, Stevenson, DW, Nakamura, K. 2012. *Zamia* (Cycadales: Zamiaceae) on Puerto Rico: Asymmetric genetic differentiation and the hypothesis of multiple introductions. *American Journal of Botany* **99:** 1828-1839.

Meerow, AW, Salas-Leiva, DE, Calonje, M, Francisco-Ortega, J, Griffith, MP, Nakamura, K, Jiménez-Rodríguez, F, Lawrus, J, Oberli, A. in press. Contrasting demographic history and population structure of *Zamia* (Cycadales: Zamiaceae) on five islands of the Greater Antilles suggests a model for population diversification in the Caribbean clade of the genus. *International Journal of Plant Sciences*.

Meerow, AWS, D. W.;Moynihan, J.;Francisco-Ortega, J. 2007. Unlocking the coontie conundrum: the potential of microsatellite DNA studies in the Caribbean *Zamia pumila* complex (Zamiaceae). *Memoirs of the New York Botanical Garden* **98:** 484-518.

Miller, MA, Pfeiffer, W, Schwartz, T. 2011. The CIPRES science gateway: a community resource for phylogenetic analyses. *Proceedings of the 2011 TeraGrid Conference: Extreme Digital Discovery.* Salt Lake City, Utah: ACM.

Montes, C, Cardona, A, McFadden, R, Moron, SE, Silva, CA, Restrepo-Moreno, S, Ramirez, DA, Hoyos, N, Wilson, J, Farris, D, Bayona, GA, Jaramillo, CA, Valencia, V, Bryan, J, Flores, JA. 2012. Evidence for middle Eocene and younger land emergence in central Panama: Implications for Isthmus closure. *Geological Society of America Bulletin* **124:** 780-799.

Morrison, DA. 2009. A framework for phylogenetic sequence alignment. *Plant Systematics and Evolution* **282:** 127-149.

Morrone, JJ. 2014. Biogeographical regionalisation of the Neotropical region. *2014* **3782:** 110.

Nagalingum, NS, Marshall, CR, Quental, TB, Rai, HS, Little, DP, Mathews, S. 2011. Recent Synchronous Radiation of a Living Fossil. *Science* **334:** 796-799.

Nichols, DJ, Johnson, KR 2009. *Plants and the K-T boundary* Cambridge, United Kingdom: Cambridge University Press.

Nicolalde-Morejón, F, Vergara-Silva, F, Gonzalez-Astorga, J, Stevenson, DW. 2010. Character-based, population-level DNA barcoding in Mexican species of *Zamia* L. (Zamiaceae: Cycadales). *Mitochondrial DNA* **21:** 51-59.

Nicolalde-Morejón, F, Vergara-Silva, F, Gonzalez-Astorga, J, Stevenson, DW, Vovides, AP, Sosa, V. 2011. A character-based approach in the Mexican cycads supports diverse multigene combinations for DNA barcoding. *Cladistics* **27:** 150-164.

Nicolalde-Morejón, F, Vovides, AP, Stevenson, DW, Sosa, V. 2008. The identity of *Zamia katzeriana* and *Z. verschaffeltii* (Zamiaceae). *Brittonia* **60:** 38-48.

Norstog, K. 1980. Chromosome numbers in *Zamia* (Cycadales). *Caryologia* **33:** 419-428.

Norstog, K. 1986. *Zamia chigua*, a case of mistaken identity? *Fairchild Tropical Garden Bulletin* **41:** 6-13.

Norstog, KJ, Nichols, TJ 1997. *The biology of the cycads* Ithaca, New York: Cornell University Press.

O'Brien, CW, Tang, W. 2015. Revision of the New World cycad weevils of the subtribe Allocorynina, with description of two new genera and three new subgenera (Coleoptera: Belidae: Oxycoryninae). *Zootaxa* **3970:** 1-87.

O'Dea, A, Lessios, HA, Coates, AG, Eytan, RI, Restrepo-Moreno, SA, Cione, AL, Collins, LS, de Queiroz, A, Farris, DW, Norris, RD, Stallard, RF, Woodburne, MO, Aguilera, O, Aubry, M-P, Berggren, WA, Budd, AF, Cozzuol, MA, Coppard, SE, Duque-Caro, H, Finnegan, S, Gasparini, GM, Grossman, EL, Johnson, KG, Keigwin, LD, Knowlton, N, Leigh, EG, Leonard-Pingel, JS, Marko, PB, Pyenson, ND, Rachello-Dolmen, PG, Soibelzon, E, Soibelzon, L, Todd, JA, Vermeij, GJ, Jackson, JBC. 2016. Formation of the Isthmus of Panama. *Science advances* **2:** 1.

Olson, K, Gorelick, R. 2011. Chromosomal fission accounts for small-scale radiations in *Zamia* (Zamiaceae; Cycadales). *Botanical Journal of the Linnean Society* **165:** 168-185.

Osborne, R, Calonje, M, Hill, KD, Stanberg, L, Stevenson, DW. 2012. The world list of cycads. *Mem. New York Bot. Gard.* **106:** 480-510.

Paradis, E, Claude, J, Strimmer, K. 2004. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **20:** 289-290.

Pérez-Farrera, MÁ, Vovides, AP, Martínez-Camilo, R, Martínez-Meléndez, N, Gómez-Domínguez, H, Galicia-Castellanos, S. 2012. *Zamia grijalvensis* sp. nov. (Zamiaceae, Cycadales) from Chiapas, Mexico with notes on hybridization and karyology. *Nordic Journal of Botany* **30:** 565-570.

Pérez-Farrera, MA, Vovides, AP, Ruiz-Castillejos, C, Galicia, S, Cibrián-Jaramillo, A, López, S. 2016. Anatomy and morphology suggest a hybrid origin of *Zamia katzeriana* (Zamiaceae). *2016* **270:** 21.

Phillips, SJ, Anderson, RP, Schapire, RE. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190:** 231-259.

Pybus, OG, Harvey, PH. 2000. Testing macro–evolutionary models using incomplete molecular phylogenies. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **267:** 2267-2272.

Rabosky, DL. 2006. LASER: A maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. *Evolutionary Bioinformatics Online* **2:** 247-250.

Rabosky, DL. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLOS ONE* **9:** e89543.

Rabosky, DL, Hurlbert, AH. 2015. Species richness at continental scales is dominated by ecological limits. *The American Naturalist* **185:** 572-583.

Rai, HS, O'Brien, HE, Reeves, PA, Olmstead, RG, Graham, SW. 2003. Inference of higher-order relationships in the cycads from a large chloroplast data set. *Molecular Phylogenetics and Evolution* **29:** 350-359.

Rambaut, A. 2012. FigTree v1. 4. *Molecular evolution, phylogenetics and epidemiology. Edinburgh, UK: University of Edinburgh, Institute of Evolutionary Biology*.

Rambaut, A, Suchard, MA, Xie, D, Drummond, AJ. 2014. Tracer v1. 6. *Available from* http://tree.bio.ed.ac.uk/software/tracer/.

Regel, E. 1876. Cycadearum, generum specierumque revisio. *Acta Hort. Petropol.* **4:** 275-320.

Revell, LJ. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3:** 217-223.

Sabato, S. 1990. West Indian and South American cycads. *Mem. New York Bot. Gard.* **57:** 173-185.

Salas-Leiva, DE, Meerow, AW, Calonje, M, Francisco-Ortega, J, Griffith, MP, Nakamura, K, Sanchez, V, Knowles, L, Knowles, D. 2017. Shifting Quaternary migration patterns in the Bahamian archipelago: Evidence from the *Zamia pumila* complex at the northern limits of the Caribbean island biodiversity hotspot. *American Journal of Botany* **104:** 757-771.

Salas-Leiva, DE, Meerow, AW, Calonje, M, Griffith, MP, Francisco-Ortega, J, Nakamura, K, Stevenson, DW, Lewis, CE, Namoff, S. 2013. Phylogeny of the cycads based on multiple single-copy nuclear genes: congruence of concatenated parsimony, likelihood and species tree inference methods. *Annals of Botany* **112:** 1263-1278.

Salas-Leiva, DE, Meerow, AW, Francisco-Ortega, J, Calonje, M, Griffith, MP, Stevenson, DW, Nakamura, K. 2014. Conserved genetic regions across angiosperms as tools to develop single-copy nuclear markers in gymnosperms: an example using cycads. *Molecular Ecology Resources* **14:** 831-845.

Sarich, VM. 1977. Rates, sample sizes, and the neutrality hypothesis for electrophoresis in evolutionary studies. *Nature* **265:** 24-28.

Schuster, J. 1932. Cycadaceae. In: Engler, AD, L. ed *Das Pflanzenreich Vol IV, part I.*

Schutzman, B. 1984. A new species of *Zamia* L. (Zamiaceae, Cycadales) from Chiapas, Mexico. *Phytologia* **55:** 299-303.

Schutzman, B, Vovides, AP. 1998. A new *Zamia* (Zamiaceae, Cycadales) from eastern Chiapas, Mexico. *Novon* **8:** 441-446.

Schutzman, B, Vovides, AP, Dehgan, B. 1988. Two New Species of *Zamia* (Zamiaceae, Cycadales) from Southern Mexico. *Botanical Gazette* **149:** 347-360.

Sclater, PL. 1858. On the general geographical distribution of the members of the Class Aves. *Journal of the Proceedings of the Linnean Society of London. Zoology* **2:** 130-136.

Smith, BT, Klicka, J. 2010. The profound influence of the Late Pliocene Panamanian uplift on the exchange, diversification, and distribution of New World birds. *Ecography* **33:** 333-342.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30:** 1312-1313.

Stevenson, D, Sabato, S, Moretti, A, De Luca, P. 1998. What is *Zamia fischeri* Miquel? *Delpinoa* **37-38:** 9-7.

Stevenson, DW. 1987. Again the West Indian zamias. *Fairchild Tropical Garden Bulletin* **42:** 23-27

Stevenson, DW. 1993. The Zamiaceae in Panama with comments on phytogeography and species relationships. *Brittonia* **45:** 1-16.

Stevenson, DW 2001. *Flora de Colombia. Monografia No. 21. Orden Cycadales* Bogota D. C., Colombia: Instituto de Ciencias Natural, Univrsidad Nacional de Colombia.

Stevenson, DW. 2004. Cycads of Colombia. *Botanical Review* **70:** 194-234.

Swofford, DL. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tang, W. 1989. Seed dispersal in the cycad *Zamia pumila* in Florida. *Canadian Journal of Botany* **67:** 2066-2070.

Tang, W. 1993. Seed dispersal and reproductive effort in *Zamia pumila*. In: Stevenson, DW, Norstog, K eds. *The biology, structure, and systematics of the cycadales. Proceedings of Cycad 90, the second international conference on cycad biology.* Milton, Queensland: Palm & Cycad Societies of Australia Ltd.

Taylor, AS, Haynes, JL, Holzman, G. 2008. Taxonomical, nomenclatural and biogeographical revelations in the *Zamia skinneri* complex of Central America (Cycadales: Zamiaceae). *Botanical Journal of the Linnean Society* **158:** 399-429.

Toledo, VM. 1982. Pleistocene changes of vegetation in tropical Mexico. In: Prance, GT ed *Biological diversification in the tropics.* New York: Columbia University Press.

Treutlein, J, Wink, M. 2002. Molecular phylogeny of cycads inferred from rbcL sequences. *Naturwissenschaften* **89:** 221-225.

van Gestel, J-P, Mann, P, Dolan, JF, Grindlay, NR. 1998. Structure and tectonics of the upper Cenozoic Puerto Rico-Virgin Islands carbonate platform as determined from seismic reflection studies. *Journal of Geophysical Research: Solid Earth* **103:** 30505-30530.

Vovides, AP. 1983. Systematic studies on the Mexican Zamiaceae. I. Chromosome numbers and karyotypes. *American Journal of Botany* **70:** 1002-1006.

Vovides, AP, Olivares, M. 1996. Karyotype polymorphism in the cycad *Zamia loddigesii* (Zamiaceae) of the Yucatan peninsula, Mexico. *Botanical Journal of the Linnean Society* **120:** 77-83.

Wallace, AR 1878. *Tropical nature, and other essays*: AMS Press.

Walters, TW, Decker Walters, DS. 1991. Patterns of allozyme diversity in the West Indies cycad *Zamia pumila* (Zamiaceae). *American Journal of Botany* **78:** 436-445.

Webb, SD. 1976. Mammalian faunal dynamics of the great American interchange. *Paleobiology* **2:** 220-234.

Weigelt, P, Steinbauer, MJ, Cabral, JS, Kreft, H. 2016. Late Quaternary climate change shapes island biodiversity. *Nature* **532:** 99.

Wendel, JF, Doyle, JJ. 1998. Phylogenetic incongruence: window into genome history and molecular evolution. In: Soltis, DE, Soltis, PS, Doyle, JJ eds. *Molecular systematics of plants II: DNA sequencing.* Boston, MA: Springer US.

Whitlock, BA, Hale, AM, Groff, PA. 2010. Intraspecific inversions pose a challenge for the *trnH-psbA* plant DNA barcode. *Plos One* **5:** e11533.

Yang, SL, Meerow, AW. 1996. The *Cycas pectinata* (Cycadaceae) complex: Genetic structure and gene flow. *International Journal of Plant Sciences* **157:** 468-483.

Yessoufou, K, Bamigboye, SO, Daru, BH, van der Bank, M. 2014. Evidence of constant diversification punctuated by a mass extinction in the African cycads. *Ecology and Evolution* **4:** 50-58.

Yu, Y, Harris, AJ, Blair, C, He, X. 2015. RASP (Reconstruct Ancestral State in Phylogenies): A tool for historical biogeography. *Molecular Phylogenetics and Evolution* **87:** 46-49.

Yu, Y, Harris, AJ, He, XJ. 2010. S-DIVA (Statistical Dispersal-Vicariance Analysis): A tool for inferring biogeographic histories. *Molecular Phylogenetics and Evolution* **56:** 848-850.

Zgurski, JM, Rai, HS, Fai, QM, Bogler, DJ, Francisco-Ortega, J, Graham, SW. 2008. How well do we understand the overall backbone of cycad phylogeny? New insights from a large, multigene plastid data set. *Molecular Phylogenetics and Evolution* **47:** 1232-1237.

Zhang, PY, Tan, HTW, Pwee, KH, Kumar, PP. 2004. Conservation of class C function of floral organ development during 300 million years of evolution from gymnosperms to angiosperms. *Plant Journal* **37:** 566-577.

Zonneveld, BJM, Lindström, AJ. 2016. Genome sizes for 71 species of *Zamia* (Cycadales: Zamiaceae) correspond with three different biogeographic regions. *Nordic Journal of Botany* **34:** 744-751.

Table 1. Locus name, description, and publication source for loci used in this study

Locus	Locus description	Reference(s)
40S	40S ribosomal protein S27-2 (RS27A)	Salas-Leiva *et al.* (2014)
CyAG	MADS-box transcription factor family AGAMOUS	Zhang *et al.* (2004), Salas-Leiva *et al.* (2013)
ATG2	ATG2, EBP1, ATEBP1, metallopeptidase M24 family protein	Salas-Leiva *et al.* (2014)
GroES	GroES-like zinc-binding alcohol dehydrogenase family protein	Salas-Leiva *et al.* (2013)
HTS	Histidyl-tRNA synthetase	Salas-Leiva *et al.* (2013)
LiSH	LisH/CRA/RING-U-box domain-containing protein	Salas-Leiva *et al.* (2014)
PEX4	PEX4, peroxin4	Salas-Leiva *et al.* (2014)
PMP22	Peroxisomal membrane 22 kDa (Mpv17/PMP22) family protein	Salas-Leiva *et al.* (2014)
WRKY4	WRKY transcription factors	Salas-Leiva *et al.* (2017)
psbK/I	Noncoding intergenic chloroplast spacer	Lahaye *et al.* (2008)

Table 2. PCR and Sequencing Primers and annealing temperatures

Locus ID	Forward primer sequence (5'-3') Name	Sequence	Reverse primer sequence (5'-3') Name	Sequence	T(°C)	PCR/SEQ
40S	40S_F2	CTGGAAAAGCGAAAGCACAA[2]	40S_R1aa	CATGACTCCAAGATAACAAG	55	PCR/SEQ
	40S_F2bm	TGGAGTTGATAGCTTGGTC	40S_R4m	CAGACGGTTGTGGTGTGT		SEQ
CyAG	CeAG_F7	CCATTTCAGAGTCCAATTCTCAG[1]	AGM_3126r	GCTTAAAGATGTGAGTGCATCTC[1]	60	PCR/SEQ
	CeAG_F8	GCAACAGGAGGCAGGAAAAC[1]	AGM3596_R	CTTAGTGCGAAGAAACTGATTCTC[1]	60	PCR/SEQ
ATG2	ATG2_F1	GTCAATAGCACTGTATGYCA[2]	ATG2_R2	ATAGGACCTTCTTGCAACA[2]	58	PCR/SEQ
GroES	GroES_F1	TYGTGGAAGCTGTGGGGA[2]	GroES_R2b	GTCTCTGGCAATTGAAGAAATG[2]	55/50*	PCR
	GroES_F1b	CTTGGAAGACAGAATTTGGCA	GroES_R2cm2	GGCAATACCMACCTGTCC		SEQ
			GroES_R2cM	CTTCAGGAAATGCCAAAT		SEQ
	GroES_F1c	CCAAGCTGATGATGGTAATTTC[2]	GroES_R2	TACATGGTCWGCTCCTAA[2]	55/50*	PCR
	CyGroES_F1cdd1	ASACATTTCYTGTTAGGGC	CyGroES_r2ab	CCACAAACCTTTCCTATC		SEQ
HTSLiSH	HTS_F1aaM	TGTTGGACAGCATGAARA	HTS_R2a	AGTCTCTCTCAACTCAAA[2]	55	PCR/SEQ
	LisH_F1	CAGGTTGAAGGCAGCATC[2]	LisH_r1m	CAGCCTCCTCTTTTGACA	58	PCR
	LisH_F1c	ATGTGGCCAGATTCATTG	LisH_R7	TTAACATCCGTCATCATACCC		SEQ
	LisH_f1am	TTTGCAGCGGCTTTCTCT	LisH_R4	GTGATTTAGTCATGAAGCAG		SEQ
PEX4	PEX4_F1	TCCAGCTAGCCATGACTGTTC[2]	PEX4_R1	GGTTTTGACCCCTATTCGGTA[2]	55	PCR/SEQ
PMP22	PMP22_F2	CAGCAGGAGATATTGGTGC[2]	PMP22_R3	AGCTWCACACACAAAGAC	60	PCR/SEQ
WRKY4	WRKY4_F2	GGTAGGAGGAAGGAAAAGAAAGT	WRKY4_R1	AATGGCTTTGGGATCATTTG	60	PCR/SEQ
psbK/I	psbK	TTAGCCTTTGTTTGGCAAG[3]	psbI	AGAGTTTGAGAGTAAGCAT[3]	52	PCR/SEQ

*Touchdown PCR; [1]Salas et al., 2013; [2]Salas et al., 2014; [3]Lahaye et al., 2008; all primers without superscript developed for this paper.

Table 3. Site models with highest posterior support selected by BMODELTEST

Locus	Posterior_support	Model (r_{ac} r_{ag} r_{at} r_{cg} r_{ct} r_{gt})
40S	33.21%	121231
CyAG	31.65%	123121
ATG2	10.84%	121321
GroES	20.53%	121343
HTS	17.69%	121321
LiSH	25.56%	123451
PEX4	10.26%	121323
PMP22	11.52%	123221
psbK/I	18.63%	123324
WRKY4	13.91%	121231

Table 4. Characters and character states for ancestral character state reconstruction using parsimony

Character	Character states
Presence of prickles on petiole	0 absent; 1 present
Prominence of teeth on leaflet margin	0 prominent; 1 not prominent;
Prominence of leaflet veins	0 prominent; 1 not prominent; 2 variable
Arborescence	0 subterranean stem; 1 arborescent stem

Table 5. Sequence characteristics of 10 loci and concatenated matrix (supermatrix) across 113 samples of *Zamia*, 1 of *Microcycas*, and 1 of *Stangeria*

Locus ID	Alignment Length	Parsimony informative characters	% Parsimony informative characters	Number trees found	Tree Length	Consistency Index (CI)	Retention Index (RI)	Homoplasy Index (HI)
*40S**	1096	91	8.3%	7970	257	0.821	0.958	0.179
CyAG	1289	111	8.6%	8950	418	0.859	0.913	0.141
ATG2	696	71	10.2%	8750	188	0.888	0.923	0.112
GroES	1991	127	6.4%	4190	450	0.871	0.947	0.129
HTS	820	63	7.7%	9400	219	0.877	0.949	0.123
*LiSH**	1263	64	5.1%	7200	231	0.853	0.965	0.147
PEX4	498	29	5.8%	3750	98	0.969	0.977	0.031
PMP22	574	39	6.8%	9750	125	0.816	0.956	0.184
WRKY4	791	48	6.1%	6220	168	0.982	0.988	0.018
PsbK/I	980	52	5.3%	9680	167	0.844	0.885	0.156
Supermatrix	9998	695	7.0%	1820	2776	0.728	0.869	0.272

* Excluding *Stangeria*

Table 6. Selected mean clade ages in millions of years (Ma) ago and their 95% highest posterior density (HPD) intervals from dated species tree analysis with *BEAST of the genus *Zamia* (fig. 3).

	Stem		Crown	
Node #	Mean	95% HPD	Mean	95% HPD
1	n/a	n/a	127.17	98.74–147.7
2	127.17	98.74–147.7	68.28	50.99–84.5
3	68.28	50.99–84.5	9.54	9–10.62
4	9.54	9–10.62	1.19	0.61–1.86
5	9.54	9–10.62	5.97	3.92–8.15
6	5.97	3.92–8.15	2.59	0.75–4.62
7	5.97	3.92–8.15	4.8	3.15–6.58
8	4.8	3.15–6.58	3.9	2.54–5.34
9	3.9	2.54–5.34	2.39	0.99–3.94
10	3.9	2.54–5.34	1.94	1.02–3.08
11	1.94	1.02–3.08	1.23	0.43–2.04
12	1.94	1.02–3.08	1.33	0.76–1.95
13	4.8	3.15–6.58	4.42	2.83–6.08
14	4.42	2.83–6.08	3.8	2.49–5.17
15	3.8	2.49–5.17	2.35	1.43–3.32
16	2.35	1.43–3.32	1.58	0.73–2.5
17	2.35	1.43–3.32	1.97	1.2–2.81
18	1.15	0.55–1.81	0.76	0.26–1.28
19	1.2	0.55–1.94	0.77	0.25–1.38
20	3.8	2.49–5.17	2.62	1.71–3.56
21	2.62	1.71–3.56	0.74	0.14–1.41
22	1.84	1.19–2.67	1.2	0.61–1.85
23	1.53	0.63–2.4	0.49	0–1.36
24	2.43	1.63–3.3	2.01	1.29–2.77
25	1.78	1.11–2.48	1.27	0.58–2.01
26	2.01	1.29–2.77	1.41	0.8–2.12

Table 7. Nodes with Dispersal and Vicariance Events from RASP DEC Analysis

Node	Ancestral Area	Child taxa	Child nodes	Event(s)	RASP Route	Probability
110	CD 100.00	*Zamia pyrophylla*	Node109: C 100.00	Vicariance: 1	CD->D|C	1
111	D 80.43 CD 19.57		Node110: CD 100.00; Node108: D 100.00	Dispersal: 1	D->D^D->CD^D->CD|D	0.8043
139	C 100.00	*Zamia obliqua*	Node138: C 100.00	Dispersal: 1	C->C^C->CD^C->CD|C	1
147	CD 100.00		Node146: C 100.00; Node131: D 97.50 CD 2.50	Vicariance: 1	CD->C|D	0.975
148	BCD 33.74 BD 33.74 BC 32.52	*Zamia soconuscensis*	Node147: CD 100.00	Vicariance: 1	BCD->B|CD	0.3374
149	B 62.63 BCD 12.79 BD 12.46 BC 12.12		Node148: BCD 33.74 BD 33.74 BC 32.52; Node104: B 100.00	Dispersal: 2	B->B^B->BCD^B->BCD|B	0.2113
151	AB 100.00		Node150: B 100.00; Node85: A 100.00	Vicariance: 1	AB->B|A	1
152	AB 52.00 A 48.00	*Microcycas calocoma*	Node151: AB 100.00	Dispersal: 1	AB->AB^A->A|AB	0.52
153	ABE 52.00 AE 48.00	*Stangeria eriopus*	Node152: AB 52.00 A 48.00	Vicariance: 1	ABE->E|AB	0.2704

Table 8. Net diversification rates of major clades of *Zamia* using different extinction rates. Rates from other studies calculated based on inferred clade ages.

		Net diversification rates		
Clade	***n***	**(e) = 0**	**(e) = 0.7**	**(e) = 0.9**
This study				
Full tree	77	0.03	0.02	0.02
Zamia stem	76	0.05	0.04	0.03
Zamia crown	75	0.39	0.32	0.22
Caribbean	9	1.26	0.88	0.45
Mainland	66	0.60	0.49	0.34
Mesoamerica	18	0.60	0.46	0.27
Fischeri	3	0.16	0.10	0.05
Central-meridional	44	0.82	0.66	0.43
Isthmus	16	0.88	0.65	0.37
South America	28	1.04	0.81	0.50
Other studies				
Gymnosperms (Crisp & Cook, 2011)	18	0.17	0.12	0.07
Zamia crown (Nagalingum et al., 2011)	31	0.32	0.27	0.19
Zamia crown (Condamine et al., 2015)	43	0.25	0.21	0.14

Table 9. Species diversity, γ values and results of the MCCR test for major clades studied. "*" indicates significant P-values where the null model of constant diversification was rejected.

Clade	Total diversity	Sampled diversity	% of diversity sampled	Gamma Statistic (γ)	Critical Value of γ (at 5% level)	P-value
Full tree	82	77	94%	11.18	-1.79	1*
Zamia stem	81	76	94%	8.67	-1.78	1*
Zamia crown	80	75	94%	1.73	-1.36	0.52
Caribbean	9	9	100%	0.97	-1.62	0.83
Mainland	71	66	93%	-0.07	-1.77	0.86
Mesoamerica	21	18	86%	0.45	-1.79	0.72
Fischeri	3	3	100%	0.60	-1.54	0.68
Central-meridional	46	44	96%	-1.24	-1.75	0.12
Isthmus	16	16	100%	-0.01	-1.63	0.50
South America	30	28	93%	-1.84	-1.75	0.04*

Table 10. Summary of variable contribution of environmental variables in MAXENT analyses

Variable	Isthmus		Mesoamerica		*Zamia soconuscensis*	
	Percent contribution	**Permutation importance**	**Percent contribution**	**Permutation importance**	**Percent contribution**	**Permutation importance**
Bio1 = Annual Mean Temperature	0.2	1.1	0.5	0.6	0.1	0
Bio2 = Mean Diurnal Range	66.6	41.2	4.1	3.6	2.1	0
Bio3 = Isothermality	4.2	6.1	10.5	6.2	0.1	0
Bio4 = Temperature Seasonality	4	14.2	36.6	48.6	0.4	0.1
Bio5 = Max Temperature of Warmest Month	0.5	1.1	0.7	0.4	2.7	0
Bio6 = Min Temperature of Coldest Month	0.4	0.3	7.1	0	0	0
Bio7 = Temperature Annual Range	5.7	6.2	6.8	4.7	0	0
Bio8 = Mean Temperature of Wettest Quarter	2.8	2.1	0.2	0.4	0	0
Bio9 = Mean Temperature of Driest Quarter	0	0	1.3	2.2	0	0
Bio10 = Mean Temperature of Warmest Quarter	0	0	0	0.1	0	0
Bio11 = Mean Temperature of Coldest Quarter	0	0.1	1.8	1	0.2	0
Bio12 = Annual Precipitation	7.8	10.5	0.7	1.1	1.6	0
Bio13 = Precipitation of Wettest Month	0	0	11.8	4.2	36.3	0.8
Bio14 = Precipitation of Driest Month	0.9	3.6	1.1	4.2	0	0
Bio15 = Precipitation Seasonality	1.8	4.1	2.8	14.3	19.1	0.6
Bio16 = Precipitation of Wettest Quarter	0.5	0.3	6.5	0.2	8.8	0.1
Bio17 = Precipitation of Driest Quarter	0.2	6.5	0.6	0.7	0	0
Bio18 = Precipitation of Warmest Quarter	1	1.9	0.6	1.5	0.7	0
Bio19 = Precipitation of Coldest Quarter	3.5	0.7	6.2	5.9	27.9	98.4

Fig. 1. Geographic distribution of the New World endemic genus *Zamia.* Occurrence data compiled from 2,594 herbarium specimen and observation records derived from first author's research database. Notice disjunct Caribbean, Mesoamerican and Central and South American geographic groupings

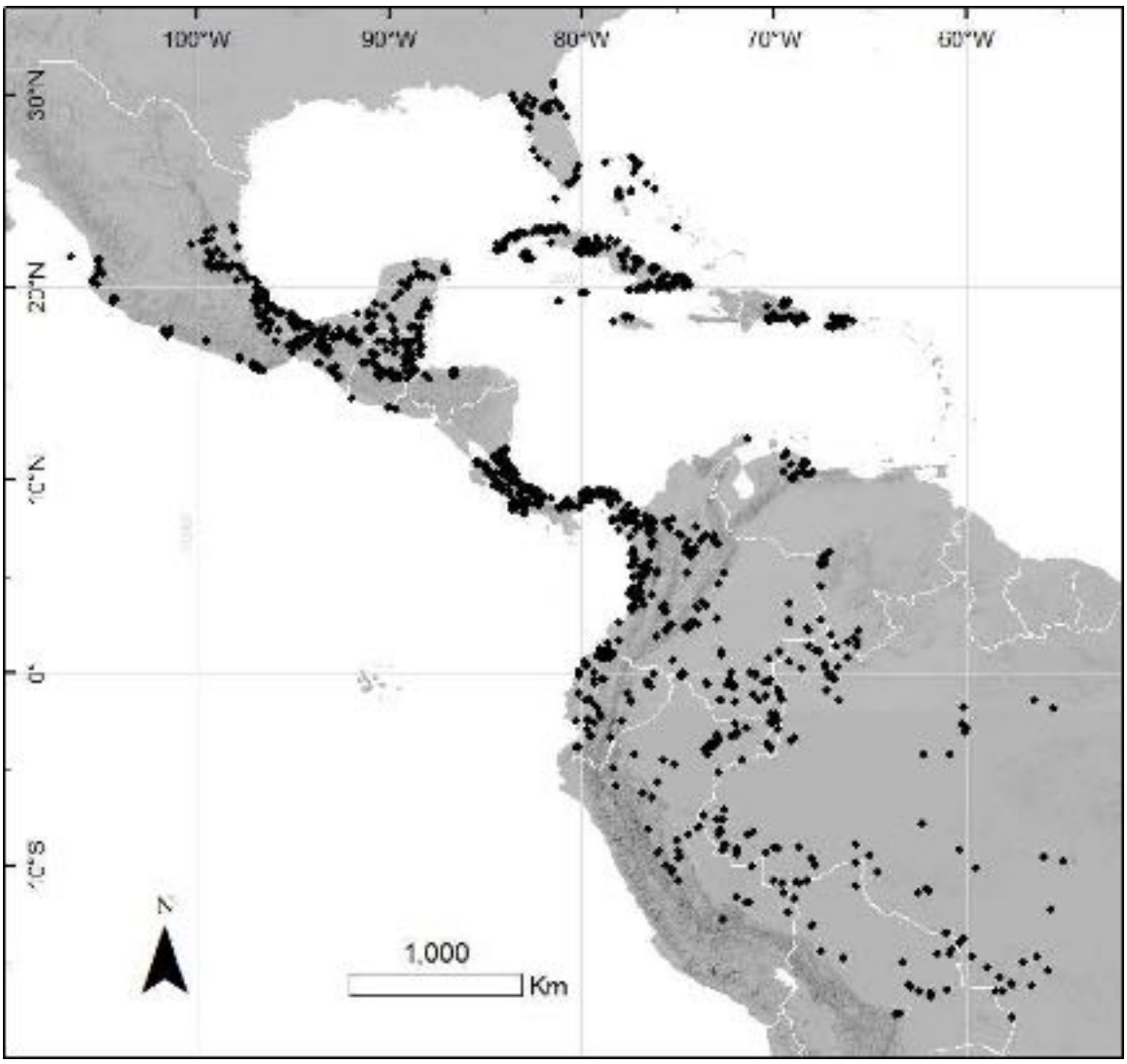


Fig. 2**.** Phylogenetic tree topologies for *Zamia* resolved by other authors. Some taxon names are revised based on current taxonomy and nomenclature, the original names shown in brackets. **(A)** Caputo *et al.* (2004): Maximum parsimony consensus tree using combined morphological data and nrDNA ITS2. Numbers at nodes indicate branch support (up to 5 steps longer trees). **(B)** Clugston *et al.* (2016): Maximum parsimony strict consensus tree found using combined nuclear (26S, CAB, NEEDLY) and plastid (matK, rbcL) genes. Numbers at nodes are Jackknife support values. **(C)** Nagalingum *et al.* (2011): Maximum likelihood tree generated from concatenated matrix combining one nuclear (PHYP) and two plastid (matK, rbcL) genes. Nodes annotated with bootstrap support values above 70%.

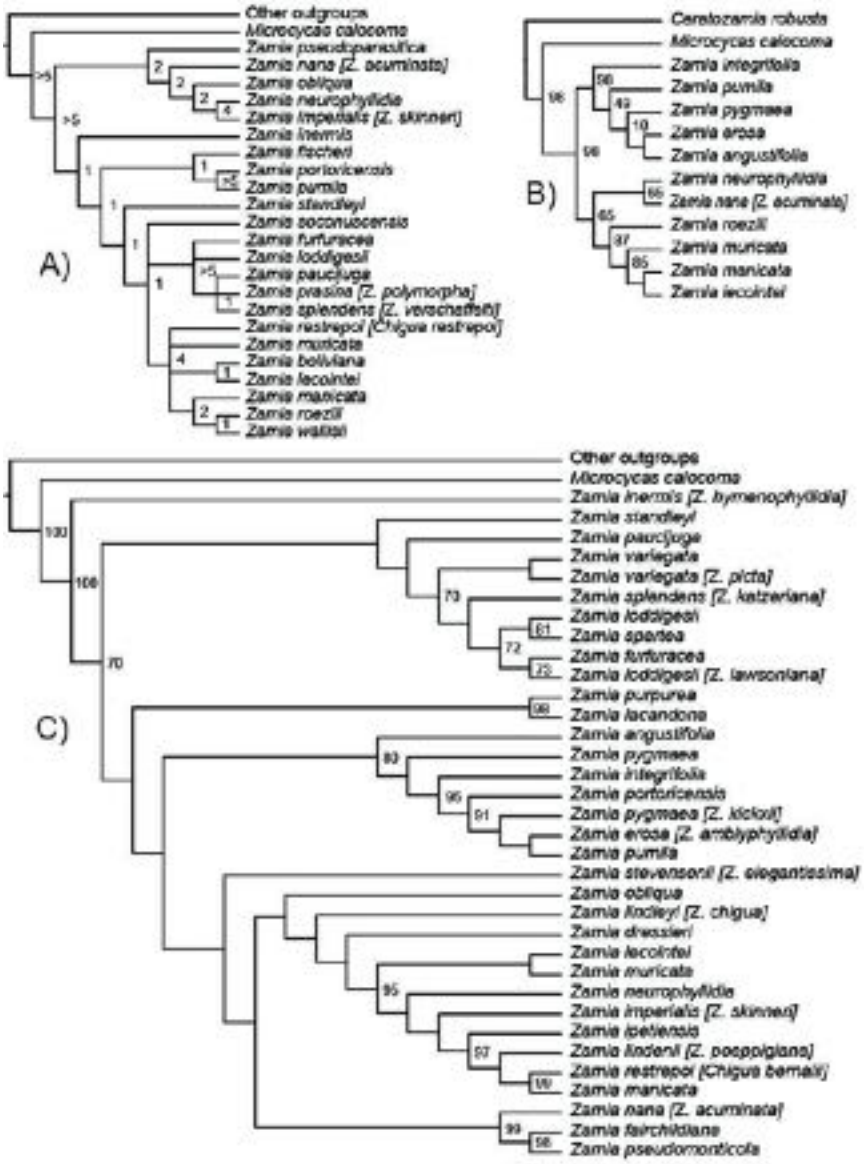


Fig. 3. Broadly congruent tree topology obtained from Maximum Parsimony (MP; PAUP), Maximum Likelihood (ML; RAXML), and Bayesian species tree (BA; BEAST) analyses of 9 SCNG loci and one plastid locus. All methods resolved the same major clades and relationships between them. Underlined numbers to the right of nodes are posterior probabilities from the BA, numbers below branches in italics are bootstrap support values for ML, and numbers above branches in **bold** are jackknife support values for the MP analysis. Numbers in parentheses next to clade names represent the number of taxa included in these clades.

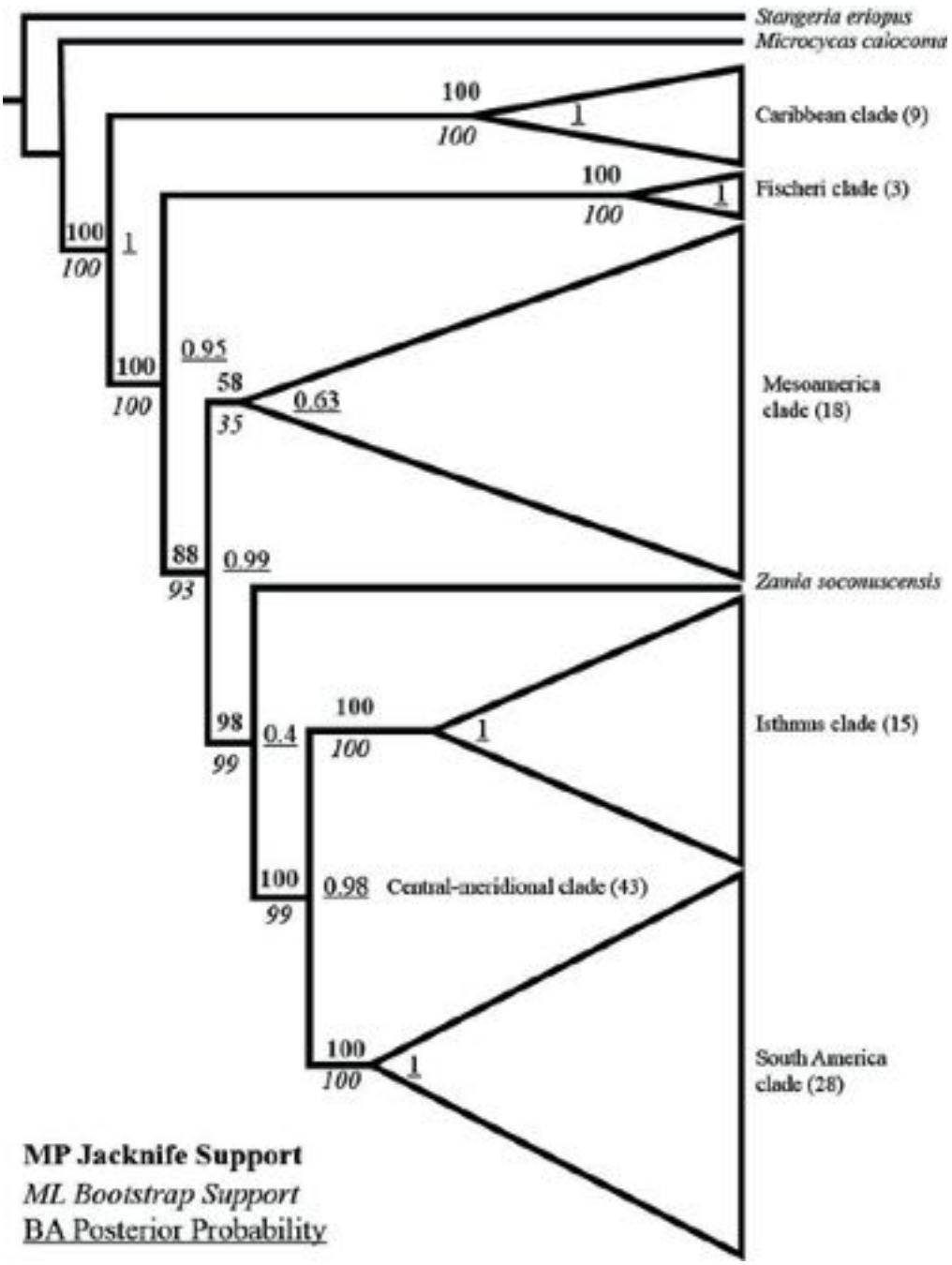


Fig. 4. Maximum clade credibility (MCC) species tree cladogram obtained from analysis in *BEAST using nine single-copy nuclear genes and one plastid locus. Posterior probabilities displayed to the right of nodes. Circles at nodes are labeled with letter codes identifying clade names with major clades color-coded. The inset map depicts the original provenance of the samples used in this study, color-coded by major clade membership.

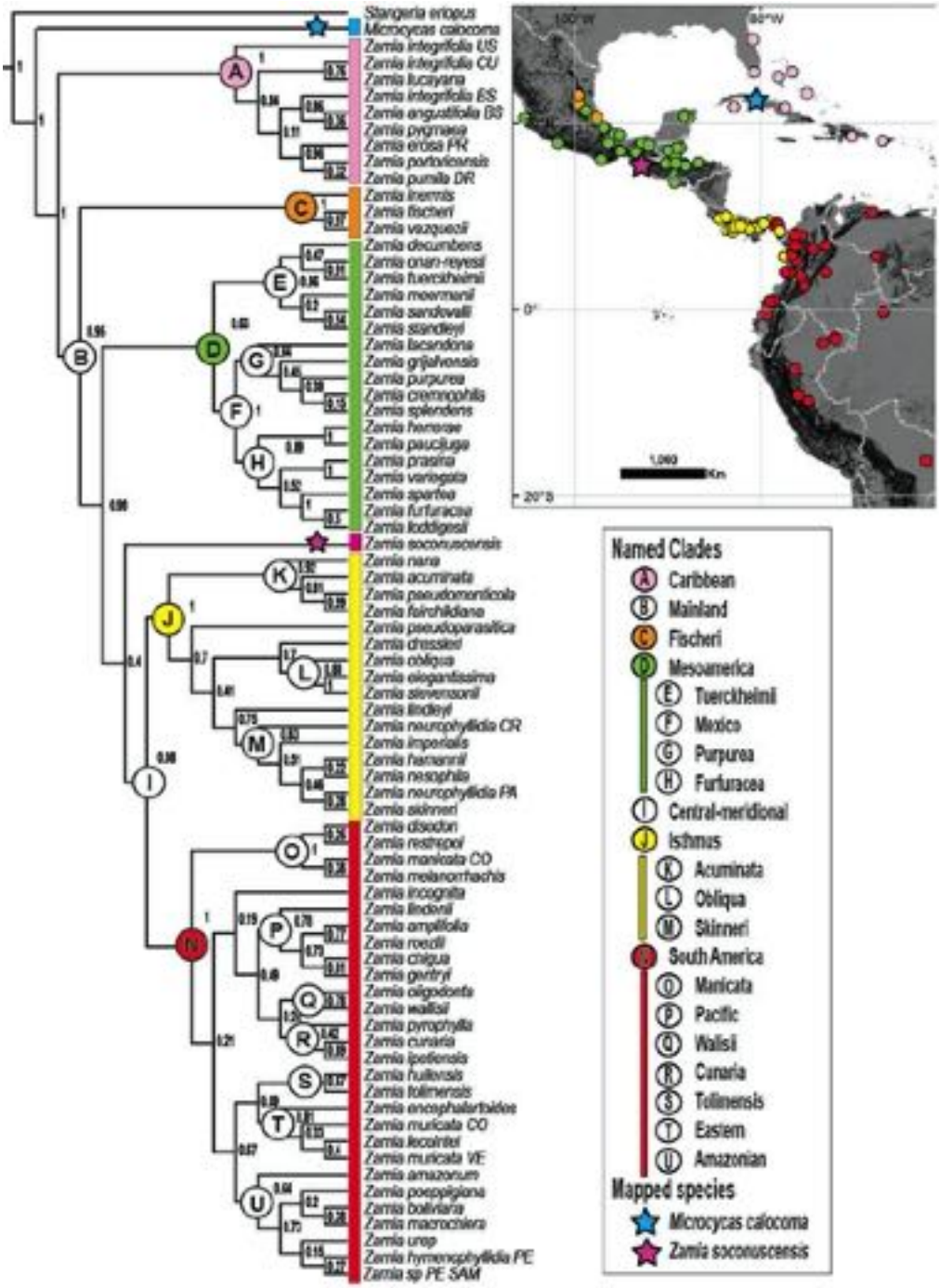


Fig. 5. Maximum clade credibility species tree chronogram obtained from analysis in *BEAST using nine single-copy nuclear genes and one plastid locus. Numbers left of nodes are mean estimated ages of clades (Ma). The gray bars represent the 95% highest posterior density (HPD) indices. The numbered nodes are referenced in (Table 6) with 95% HPD indices for stem and crown nodes. Nodes 1 and 2 (connected to dotted branches) are out of range of main figure, their true positions and ages are shown in the inset. Holocene epoch not shown. Posterior probability support values for all nodes reported in fig. 4.

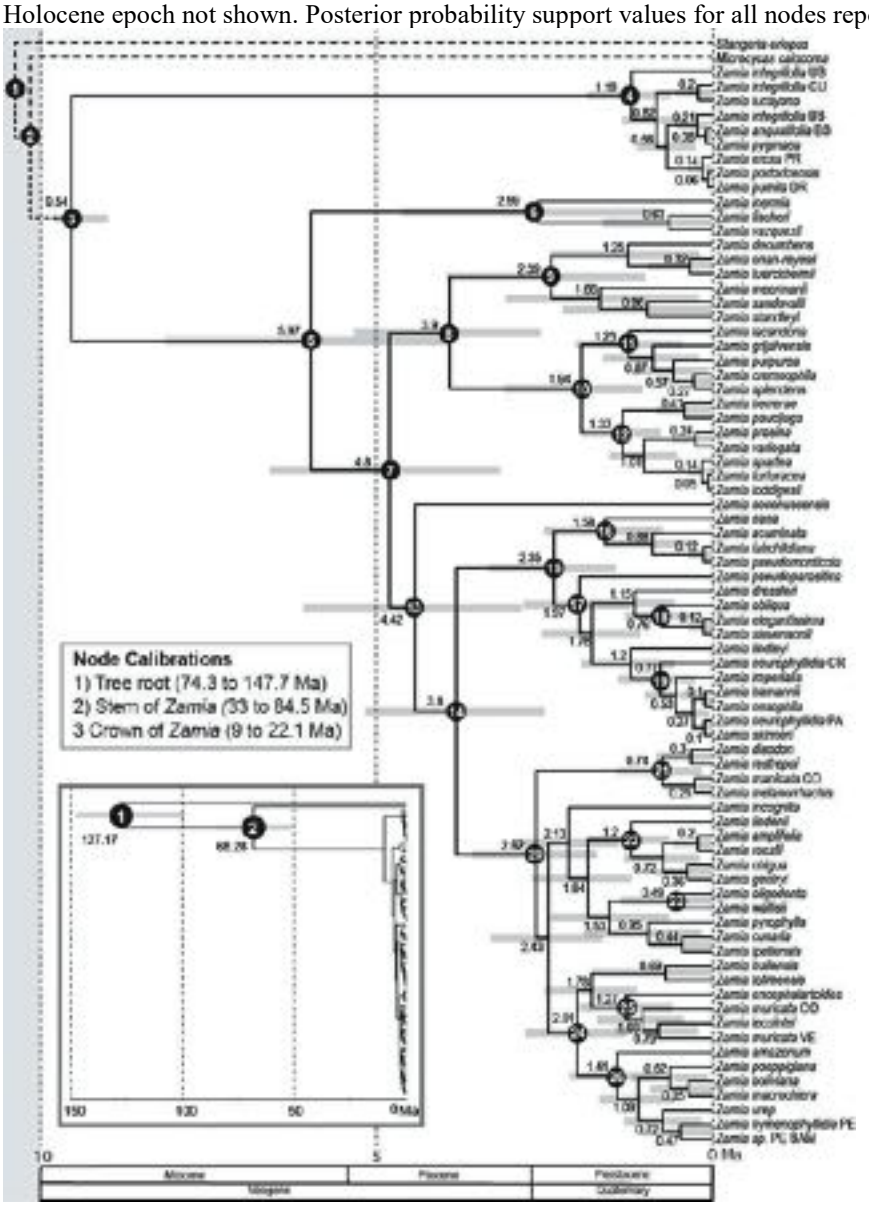


Fig. 6. Per-million-year diversification rates and γ values for major *Zamia* clades. Only the γ value for the South America clade (*) was significant, indicating diversification was initially high and decreased over time. The γ value for all other clades was not significant, indicating a constant diversification rate.

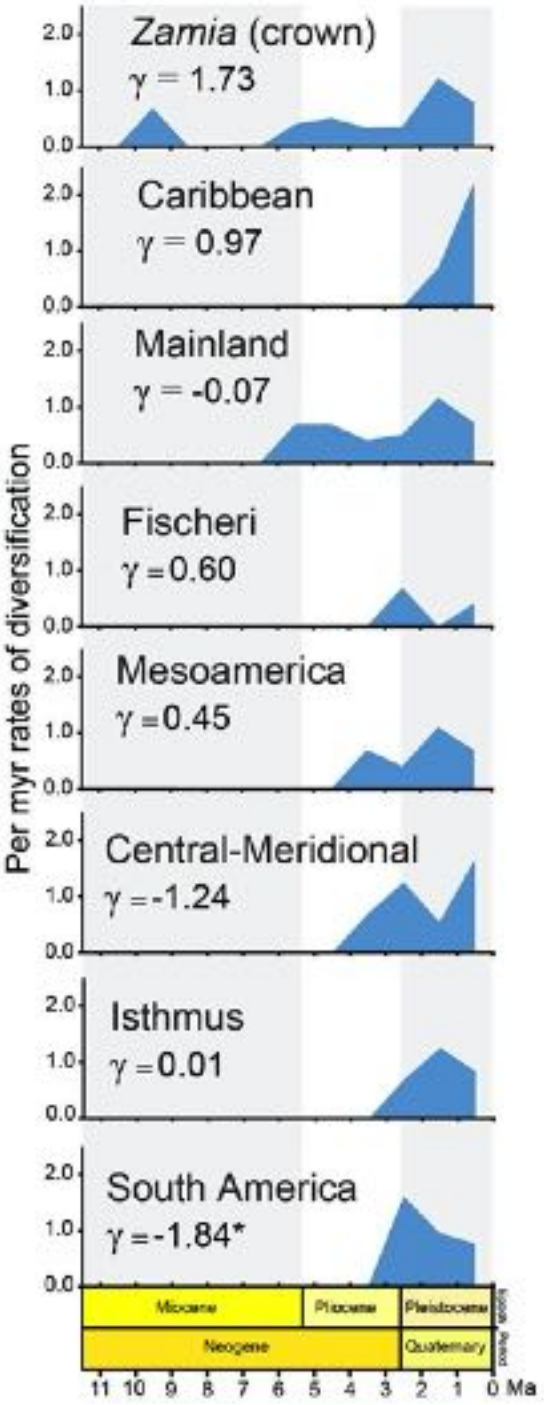


Fig. 7. Historical biogeography of *Zamia* inferred using a statistical dispersal-vicariance analysis (S-DIVA) as implemented in RASPusing the MCC chronogram tree shown in figure 4. The proportions of colors in circles represent possible ancestral geographic ranges at each node, and the rings around the circles represent dispersal and/or vicariance events. The two basal-most nodes are out of scale, their mean ages indicated next to the asterisks. The inset map shows the geographic distribution of the *Zamia* samples color-coded by their area assignations.

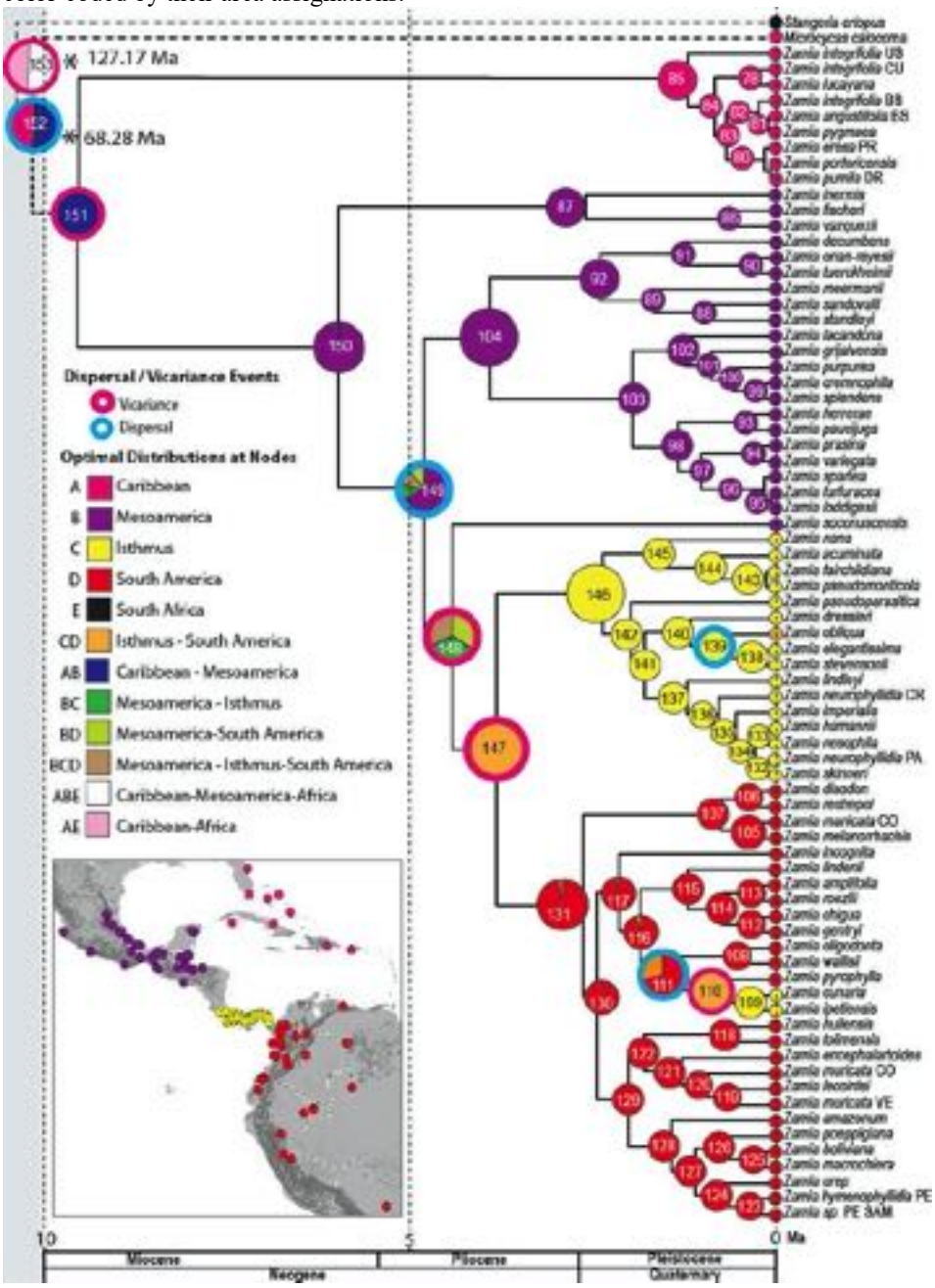


Fig. 8. Disjunct distribution of mainland *Zamia* in Central America separates the Mesoamerica and Isthmus clades. Niche modeling performed on each clade as well as *Zamia soconuscensis* with MaxEnt software using Chelsa climate layers and 767 occurrence records suggests that dispersal into the intervening region may be limited by unsuitable habitat. A) Occurrence records used in analysis; B) Close-up of distribution gap; C) Habitat suitability of Mesoamerica clade; D) Habitat suitability of Isthmus clade; E) Ocurrence records for *Z. soconuscensis*; F) Habitat suitability for *Z. soconuscensis*.

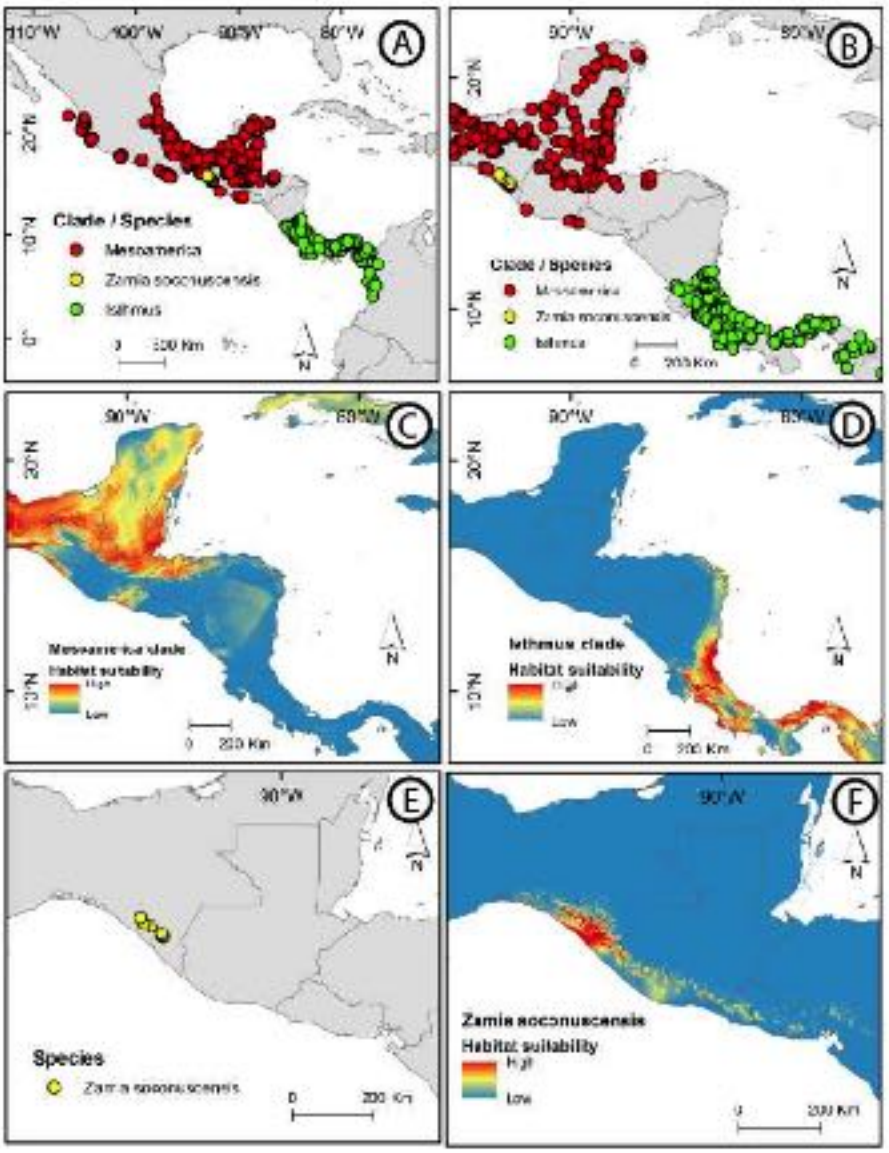
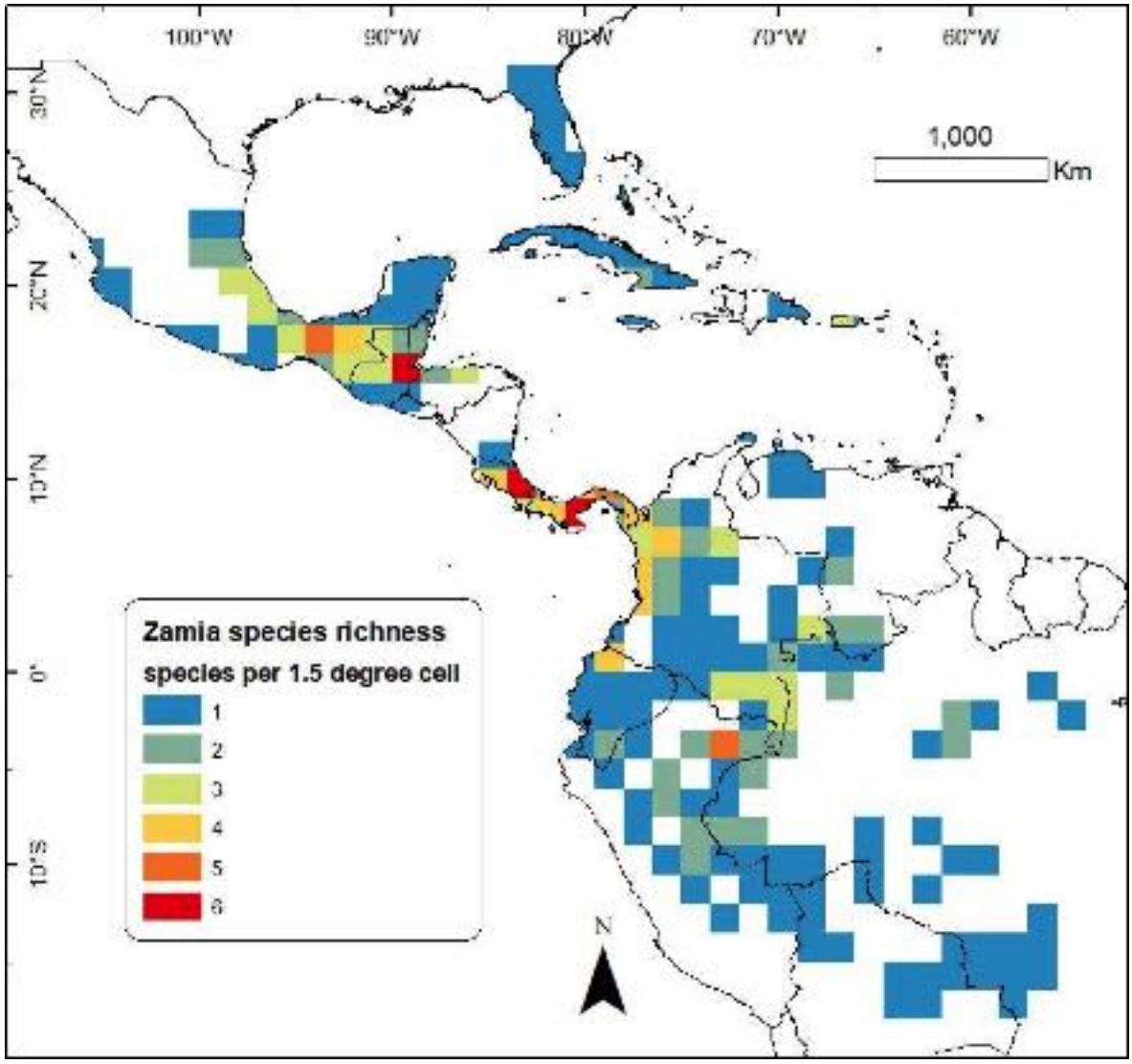


Fig. 9. Species richness of *Zamia*



APPENDICES

APPENDIX A: ACCESSIONS SAMPLED FOR MOLECULAR ANALYSES

The following information is provided for each sample: Species name and authorship, sample ID, sample provenance or pedigree (country and state or province), representative herbarium specimen (collector, herbarium repository), DNA sample source (source and botanical garden accession numbers), Genbank accession numbers (from left to right, dashes indicate no sequence available: 40S, ATG2, CyAg, GroES, HTS, LiSH, PEX4, PMP22, psbK/I, WRKY4. Herbarium acronyms follow Index Herbariorum (http://sweetgum.nybg.org/science/ih/). Abbreviations used for DNA sample sources: FTBG: Fairchild Tropical Botanic Garden, HS: Herbarium specimen, JBC: Jardín Botánico Francisco Javier Clavijero, Xalapa, Veracruz, Mexico, LA: Harold Lyon Arboretum, Manoa, Hawaii, MBC: Montgomery Botanical Center, Coral Gables, FL, U.S.A., MS: Marie Selby Botanical Gardens, Sarasota, Florida, U.S.A., NN: Nong Nooch Tropical Botanical Garden, Pattaya, Thailand, USDA-ARS: United States Department of Agriculture, Agricultural Resource Service (DNA repository), Miami, FL, U.S.A.

***Microcycas calocoma* (Miq.) A.DC.,** *Microcycas calocoma*, Cuba (Pinar del Rio), Calonje MBC12-006 (FTG), MBC 59875*B, MG148379, MG148494, MG148953, MG148609, MG148724, MG148838, MG149041, MG149183, MG149271, MG149413; ***Stangeria eriopus* (Kunze) Baill.,** *Stangeria eriopus*, South Africa, Calonje MBC12-001(FTG), MBC 9868*C, -, MG148505, MG148964, MG148620, MG148735, -, MG149079, MG149194, MG149298, MG149424; ***Zamia acuminata* Oerst. ex Dyer,** *Zamia acuminata* CR1, Costa Rica (San José), Calonje et al. MAC04-001(1) (CR, FTG), MBC 20041004, MG148390, MG148516, MG148975, MG148631, MG148746, MG148849, MG149090, MG149205, MG149309, MG149435; ***Zamia acuminata* Oerst. ex Dyer,** *Zamia acuminata* CR2, Costa Rica (San José), Calonje et al. MAC04-001(1) (CR, FTG), MBC 20041004, MG148353, MG148527, MG148986, MG148642, MG148757, MG148860, MG149101, MG149216, MG149320, MG149446; ***Zamia amazonum* D.W.Stev.,** *Zamia amazonum* BR, Brazil (Amazonas), Madison & Kennedy 6563 (F, SEL, US), LA L-78.1374, MG148401, MG148538, MG148997, MG148653, MG148768, MG148871, MG149112, MG149227, MG149331, MG149457; ***Zamia amazonum* D.W.Stev.,** *Zamia amazonum* PE,

Peru (Loreto), Lindstrom s.n. (FTG, L, V), NN, MG148412, MG148549, MG149008, MG148664, MG148779, MG148882, MG149123, MG149238, MG149342, MG149468; ***Zamia amplifolia* W.Bull ex Mast.,** *Zamia amplifolia*, Colombia (Valle del Cauca), Bogler 1223 (FTG), FTBG 91596, MG148423, MG148560, MG149019, MG148675, MG148790, MG148893, MG149134, MG149249, MG149353, MG149479; ***Zamia angustifolia* Jacq.,** *Zamia angustifolia* BS, Bahamas (Eleuthera), Calonje et al. BS10-200 (FTG, NY), USDA-ARS ZBE103, MG148434, MG148571, MG149030, MG148686, MG148801, MG148904, MG149145, MG149260, MG149364, MG149490; ***Zamia boliviana* (Brongn.) A.DC.,** *Zamia boliviana*, Bolivia, Little & Stevenson 1067 (FTG), MBC 981682*A, MG148445, MG148468, MG148927, MG148583, MG148698, MG148915, MG149042, MG149157, MG149375, MG149387; ***Zamia chigua* Seem.,** *Zamia chigua* 1, Colombia (Chocó), Wilder et al. s.n. (F), LA L-81.0783, MG148456, MG148479, MG148938, MG148594, MG148709, MG148813, MG149053, MG149168, MG149272, MG149398; ***Zamia chigua* Seem.,** *Zamia chigua* 2, Colombia (Valle del Cauca), Little & Stevenson 1068 (FTG), MBC 20010802*A, MG148354, MG148486, MG148945, MG148601, MG148716, MG148824, MG149060, MG149175, MG149283, MG149405; ***Zamia cremnophila* Vovides, Schutzman & Dehgan,** *Zamia cremnophila*, Mexico (Tabasco), Bogler 1225 (FTG), FTBG 87339C, MG148365, MG148487, MG148946, MG148602, MG148717, MG148830, MG149061, MG149176, MG149290, MG149406; ***Zamia cunaria* Dressler & D.W.Stev.,** *Zamia cunaria* 1, Panama (Panamá), Stevenson 1201 (FTG, NY, PMA, U), MBC 2000269*A, MG148371, MG148488, MG148947, MG148603, MG148718, MG148831, MG149062, MG149177, MG149291, MG149407; ***Zamia cunaria* Dressler & D.W.Stev.,** *Zamia cunaria* 2, Panama (Panamá), Stevenson 1201 (FTG, NY, PMA, U), MBC 2000311*A, MG148372, MG148489, MG148948, MG148604, MG148719, MG148832, MG149063, MG149178, MG149292, MG149408; ***Zamia decumbens* Calonje, Meerman, M.P.Griff. & Hoese,** *Zamia decumbens* 1, Belize (Toledo), Calonje et al. BZ08-201 (BRH, FTG, MO, NY, XAL), MBC 20080715, MG148373, MG148490, MG148949, MG148605, MG148720, MG148833, MG149064, MG149179, MG149293, MG149409; ***Zamia decumbens* Calonje, Meerman, M.P.Griff. & Hoese,** *Zamia decumbens* 2, Belize (Toledo), Calonje et al. BZ08-180 (FTG), MBC 20080676, MG148374, MG148491, MG148950, MG148606, MG148721, MG148834, MG149065, MG149180, MG149294, MG149410; ***Zamia disodon* D.W.Stev. & Sabato,**

Zamia disodon, Colombia (Antioquia), Lindstrom s.n. (FTG, L, V), NN, MG148375, MG148492, MG148951, MG148607, MG148722, MG148835, MG149066, MG149181, MG149295, MG149411; ***Zamia dressleri* D.W.Stev.,** *Zamia dressleri* 1, Panama (Colón), Stevenson et al. 1145 (FTG), MBC 2000332*A, MG148376, MG148493, MG148952, MG148608, MG148723, MG148836, MG149067, MG149182, MG149296, MG149412; ***Zamia dressleri* D.W.Stev.,** *Zamia dressleri* 2, Panama (Kuna Yala), De Nevers, G. 6565 (STRI), SEL 1996-0009, MG148377, MG148495, MG148954, MG148610, MG148725, MG148837, MG149069, MG149184, MG149297, MG149414; ***Zamia elegantissima* Schutzman, Vovides & R.S.Adams,** *Zamia elegantissima* 1, Panama (Colón), Little & Stevenson 1149 (FTG), MBC 2000775*G, MG148378, MG148496, MG148955, MG148611, MG148726, MG148839, MG149070, MG149185, MG149299, MG149415; ***Zamia elegantissima* Schutzman, Vovides & R.S.Adams,** *Zamia elegantissima* 2, Panama (Kuna Yala), De Nevers & Henderson 6320 (MO), FTBG 93558*B, MG148380, MG148497, MG148956, MG148612, MG148727, MG148840, MG149071, MG149186, MG149300, MG149416; ***Zamia encephalartoides* D.W.Stev.,** *Zamia encephalartoides*, Colombia (Santander), Espinosa 2011-003 (FTG), MBC 94910*A, MG148381, MG148498, MG148957, MG148613, MG148728, MG148841, MG149072, MG149187, MG149301, MG149417; ***Zamia erosa* O.F.Cook & G.N.Collins,** *Zamia erosa* PR, Puerto Rico (Quebradillas), Meerow and Ayala-Silva 3100, USDA-ARS ZPR413, MG148382, MG148499, MG148958, MG148614, MG148729, MG148842, MG149073, MG149188, MG149302, MG149418; ***Zamia fairchildiana* L.D.Gómez,** *Zamia fairchildiana* CR, Costa Rica (Puntarenas), Liesner 2860 (CR, MO), MBC 20070835, MG148383, MG148500, MG148959, MG148615, MG148730, MG148843, MG149074, MG149189, MG149303, MG149419; ***Zamia fairchildiana* L.D.Gómez,** *Zamia fairchildiana* PA, Panama (Chiriquí), Michael Calonje MAC08-005 (PMA), MBC 20080102, MG148384, MG148501, MG148960, MG148616, MG148731, MG148844, MG149075, MG149190, MG149304, MG149420; ***Zamia fischeri* Miq.,** *Zamia fischeri* MX SLP, Mexico (San Luis Potosí), Nicolalde-Morejon 1620 (XAL), JBC 2008-089B, MG148385, MG148502, MG148961, MG148617, MG148732, MG148845, MG149076, MG149191, MG149305, MG149421; ***Zamia fischeri* Miq.,** *Zamia fischeri* MX SLP2, Mexico (San Luis Potosí), Calonje FTG18-01 (FTG), FTBG 76845*A, MG148386, MG148503, MG148962, MG148618, MG148733, MG148846, MG149077, MG149192,

MG149306, MG149422; ***Zamia fischeri* Miq.,** *Zamia fischeri* MX TAM, Mexico (Tamaulipas), Little & Stevenson 1070 (FTG), MBC 20010205, MG148387, MG148504, MG148963, MG148619, MG148734, MG148847, MG149078, MG149193, MG149307, MG149423; ***Zamia furfuracea* L.f.,** *Zamia furfuracea* 1, Mexico (Veracruz), Little & Stevenson 1098 (FTG), MBC 20010214*H, MG148388, MG148506, MG148965, MG148621, MG148736, MG148848, MG149080, MG149195, MG149308, MG149425; ***Zamia furfuracea* L.f.,** *Zamia furfuracea* 2, Mexico (Veracruz), Vovides 567 (XAL), MBC 99795*C, MG148389, MG148507, MG148966, MG148622, MG148737, MG148850, MG149081, MG149196, MG149310, MG149426; ***Zamia gentryi* Dodson,** *Zamia gentryi*, Ecuador (Esmeraldas), Luther et al. 1235 (SEL), MS 1995-0429 A, MG148391, MG148508, MG148967, MG148623, MG148738, MG148851, MG149082, MG149197, MG149311, MG149427; ***Zamia grijalvensis* Pérez-Farr., Vovides & Mart.-Camilo,** *Zamia grijalvensis*, Mexico (Chiapas), N. Martínez 246 (XAL), JBC 2009-018A, MG148392, MG148509, MG148968, MG148624, MG148739, MG148852, MG149083, MG149198, MG149312, MG149428; ***Zamia hamannii* A.S.Taylor, J.L.Haynes & Holzman,** *Zamia hamannii*, Panama (Bocas del Toro), Taylor et al. ASTB04-ZE1 (PMA), MBC 20040867*C, MG148393, MG148510, MG148969, MG148625, MG148740, MG148853, MG149084, MG149199, MG149313, MG149429; ***Zamia herrerae* S.Calderón & Standl.,** *Zamia herrerae* MX, Mexico (Chiapas), Calonje MBC18-01 (FTG), MBC 20010256*B, MG148394, MG148561, MG148970, MG148676, MG148791, MG148903, MG149135, MG149250, MG149365, MG149480; ***Zamia herrerae* S.Calderón & Standl.,** *Zamia herrerae* SV, El Salvador (Sonsonate), Bogler 1186 (FTG), FTBG 93423B, MG148395, MG148559, MG148971, MG148674, MG148789, MG148902, MG149133, MG149248, MG149363, MG149478; ***Zamia huilensis* Calonje, H.E.Esquivel & D.W.Stev.,** *Zamia huilensis*, Colombia (Huila), Little 9273 (US), HS, MG148396, MG148511, MG148972, MG148626, MG148741, MG148854, MG149085, MG149200, MG149314, MG149430; ***Zamia hymenophyllidia* D.W.Stev.,** *Zamia hymenophyllidia* PE, Peru (Loreto), Lindstrom s.n. (FTG, L, V), NN 16022A, MG148397, MG148512, MG148973, MG148627, MG148742, MG148855, MG149086, MG149201, MG149315, MG149431; ***Zamia imperialis* A.S.Taylor, J.L.Haynes & Holzman,** *Zamia imperialis* 1, Panama (Coclé), Taylor & Quirós ASTB01-LRica1 (PMA), MBC 2000401*A, MG148398, MG148513, MG148974, MG148628, MG148743, MG148856, MG149087,

MG149202, MG149316, MG149432; ***Zamia imperialis*** **A.S.Taylor, J.L.Haynes & Holzman,** *Zamia imperialis* 2, Panama (Veraguas), Espinosa 2011-005 (FTG), MBC 2000266*A, MG148399, MG148514, MG148926, MG148629, MG148744, MG148857, MG149088, MG149203, MG149317, MG149433; ***Zamia incognita*** **A.Lindstr. & Idárraga,** *Zamia incognita*, Colombia (Antioquia), Lindstrom s.n. (FTG, L, V), NN, MG148400, MG148515, MG148976, MG148630, MG148745, MG148858, MG149089, MG149204, MG149318, MG149434; ***Zamia inermis*** **Vovides, J.D.Rees & Vázq.Torres,** *Zamia inermis*, Mexico (Veracruz), Espinosa 2011-006 (FTG), MBC 92143*D, MG148402, MG148517, MG148977, MG148632, MG148747, MG148859, MG149091, MG149206, MG149319, MG149436; ***Zamia integrifolia*** **L.f.,** *Zamia integrifolia* BS, Bahamas (Andros), Calonje & Meerow BS10-030.(FTG), USDA-ARS ZBA104, MG148403, MG148518, MG148978, MG148633, MG148748, MG148861, MG149092, MG149207, MG149321, MG149437; ***Zamia integrifolia*** **L.f.,** *Zamia integrifolia* CU, Cuba (Camaguey), Espinosa 2011-001 (FTG), MBC 9753*A, MG148404, MG148467, MG148979, MG148582, MG148697, MG148812, MG149068, MG149156, MG149322, MG149386; ***Zamia integrifolia*** **L.f.,** *Zamia integrifolia* US1, USA (Florida), Haynes JLH06-015 (FTG), MBC 20060452*A, MG148405, MG148519, MG148980, MG148634, MG148749, MG148862, MG149093, MG149208, MG149323, MG149438; ***Zamia integrifolia*** **L.f.,** *Zamia integrifolia* US2, USA (Florida), Haynes JLH05-284 (FTG), MBC 20050880*B, MG148406, MG148520, MG148981, MG148635, MG148750, MG148863, MG149094, MG149209, MG149324, MG149439; ***Zamia ipetiensis*** **D.W.Stev.,** *Zamia ipetiensis* 1, Panama (Kuna de Madungandí), Stevenson & Valdespinos 1159 (FTG, MO, NY, U), FTBG 89166, MG148407, MG148521, MG148982, MG148636, MG148751, MG148864, MG149095, MG149210, MG149325, MG149440; ***Zamia ipetiensis*** **D.W.Stev.,** *Zamia ipetiensis* 2, Panama (Kuna Yala), Calonje et al. MAC07-005 (PMA), MBC 20070592, MG148408, MG148522, MG148983, MG148637, MG148752, MG148865, MG149096, MG149211, MG149326, MG149441; ***Zamia ipetiensis*** **D.W.Stev.,** *Zamia ipetiensis* 3, Panama (Kuna de Wargandí), Calonje et al. MAC08-086 (PMA), MBC 20080111, MG148409, MG148523, MG148984, MG148638, MG148753, MG148866, MG149097, MG149212, MG149327, MG149442; ***Zamia lacandona*** **Schutzman & Vovides,** *Zamia lacandona*, Mexico (Chiapas), Bogler 1227 (FTG), FTBG 93932, MG148410, MG148526, MG148985, MG148641, MG148756, MG148869, MG149100, MG149215, MG149330,

MG149445; ***Zamia lecointei* Ducke,** *Zamia lecointei* 1, Venezuela (Puerto Ayacucho), Bogler 1191 (FTG), Stevenson 1143 (U), FTBG 88497, MG148411, MG148528, MG148987, MG148643, MG148758, MG148870, MG149102, MG149217, MG149332, MG149447; ***Zamia lecointei* Ducke,** *Zamia lecointei* 2, Venezuela (Puerto Ayacucho), Bogler 1191 (FTG), Stevenson 1143 (U), FTBG 88500D, MG148413, MG148529, MG148988, MG148644, MG148759, MG148872, MG149103, MG149218, MG149333, MG149448; ***Zamia lindenii* Regel ex André,** *Zamia lindenii* 1, Ecuador (Esmeraldas), Espinosa 2011-007 (FTG), MBC 20001000*F, MG148414, MG148530, MG148989, MG148645, MG148760, MG148873, MG149104, MG149219, MG149334, MG149449; ***Zamia lindenii* Regel ex André,** *Zamia lindenii* 2, Ecuador (Los Rios), Gentry & Dodson 18045 (AUU, MO), FTBG 90351, MG148415, MG148531, MG148990, MG148646, MG148761, MG148874, MG149105, MG149220, MG149335, MG149450; ***Zamia lindleyi* Warsz. ex A.Dietr.,** *Zamia lindleyi*, Panama (Chiriquí), Calonje et al. PA08-123 (PMA), MBC 20010802*A, MG148416, MG148532, MG148991, MG148647, MG148762, MG148875, MG149106, MG149221, MG149336, MG149451; ***Zamia loddigesii* Miq.,** *Zamia loddigesii* MX CHP, Mexico (Chiapas), Espinosa 2011-002 (FTG), MBC 99801*A, MG148417, MG148533, MG148992, MG148648, MG148763, MG148876, MG149107, MG149222, MG149337, MG149452; ***Zamia loddigesii* Miq.,** *Zamia loddigesii* MX OAX, Mexico (Oaxaca), Walters TW-30-05 (XAL), MBC 93948*D, MG148418, MG148534, MG148993, MG148649, MG148764, MG148877, MG149108, MG149223, MG149338, MG149453; ***Zamia loddigesii* Miq.,** *Zamia loddigesii* MX PUE, Mexico (Puebla), Calonje MBC18-02, MBC 99797*K, MG148419, MG148535, MG148994, MG148650, MG148765, MG148878, MG149109, MG149224, MG149339, MG149454; ***Zamia loddigesii* Miq.,** *Zamia loddigesii* MX SLP, Mexico (San Luis Potosí), -, MBC 2000797*A, MG148420, MG148536, MG148995, MG148651, MG148766, MG148879, MG149110, MG149225, MG149340, MG149455; ***Zamia loddigesii* Miq.,** *Zamia loddigesii* MX TAB, Mexico (Tabasco), A. Vovides 1347 (XAL), JBC 2000-036.01, MG148421, MG148537, MG148996, MG148652, MG148767, MG148880, MG149111, MG149226, MG149341, MG149456; ***Zamia loddigesii* Miq.,** *Zamia loddigesii* MX VER, Mexico (Veracruz), Stevenson et al. 538 (FTG), MBC 83311*B, MG148422, MG148539, MG148998, MG148654, MG148769, MG148881, MG149113, MG149228, MG149343, MG149458; ***Zamia lucayana* Britton,** *Zamia lucayana*, Bahamas

(Long Island), Calonje et al. BS09-036 (FTG), USDA-ARS ZBLI228, MG148424, MG148540, MG148999, MG148655, MG148770, MG148883, MG149114, MG149229, MG149344, MG149459; ***Zamia macrochiera* D.W.Stev.,** *Zamia macrochiera*, Peru (Loreto), Plowman 7275 (F, GH), LA L-82.0862, MG148425, MG148541, MG149000, MG148656, MG148771, MG148884, MG149115, MG149230, MG149345, MG149460; ***Zamia manicata* Linden ex Regel,** *Zamia manicata* CO, Colombia (Antioquia), Stevenson et al. 604 (FTG, HUA), FTBG 84272*J, MG148426, MG148542, MG149001, MG148657, MG148772, MG148885, MG149116, MG149231, MG149346, MG149461; ***Zamia meermanii* Calonje,** *Zamia meermanii*, Belize (Belize), Calonje et al BZ08-152 (BRH, FTG, MO, NY, XAL), MBC 20080671, MG148427, MG148543, MG149002, MG148658, MG148773, MG148886, MG149117, MG149232, MG149347, MG149462; ***Zamia melanorrhachis* D.W.Stev.,** *Zamia melanorrhachis*, Colombia, Lindstrom s.n. (FTG, L, V), NN 13364B, MG148428, MG148544, MG149003, MG148659, MG148774, MG148887, MG149118, MG149233, MG149348, MG149463; ***Zamia muricata* Willd.,** *Zamia muricata* CO, Colombia (Meta), Lindstrom s.n. (L, V), NN, MG148429, MG148547, MG149004, MG148662, MG148777, MG148890, MG149121, MG149236, MG149351, MG149466; ***Zamia muricata* Willd.,** *Zamia muricata* VE1, Venezuela (Carabobo), H.E. Luther s.n. (SEL 075089), SEL 1995-0433R, MG148431, MG148545, MG149006, MG148660, MG148775, MG148888, MG149119, MG149234, MG149349, MG149464; ***Zamia muricata* Willd.,** *Zamia muricata* VE2, Venezuela (Carabobo), Bogler 1226 (FTG), FTBG 88495, MG148430, MG148546, MG149005, MG148661, MG148776, MG148889, MG149120, MG149235, MG149350, MG149465; ***Zamia nana* A.Lindstr., Calonje, D.W.Stev. & A.S.Taylor,** *Zamia nana* 1, Panama (Coclé), Bogler 1215 (FTG), FTBG 89159, MG148432, MG148548, MG149007, MG148663, MG148778, MG148891, MG149122, MG149237, MG149352, MG149467; ***Zamia nana* A.Lindstr., Calonje, D.W.Stev. & A.S.Taylor,** *Zamia nana* 2, Panama (Coclé), Stevenson & Valdespinos 1147 (FTG, PMA, MO), MBC 20020234*B, MG148433, MG148550, MG149009, MG148665, MG148780, MG148892, MG149124, MG149239, MG149354, MG149469; ***Zamia nesophila* A.S.Taylor, J.L.Haynes & Holzman,** *Zamia nesophila*, Panama (Bocas del Toro), N. Espinosa 2011-008 (FTG), MBC 20010123*A, MG148435, MG148551, MG149010, MG148666, MG148781, MG148894, MG149125, MG149240, MG149355, MG149470; ***Zamia neurophyllidia* D.W.Stev.,** *Zamia neurophyllidia*

CR, Costa Rica, Burger 3877 (US), FTBG 91188, MG148436, MG148552, MG149011, MG148667, MG148782, MG148895, MG149126, MG149241, MG149356, MG149471; ***Zamia neurophyllidia* D.W.Stev.,** *Zamia neurophyllidia* PA, Panama (Bocas del Toro), Correa et al 3078 (PMA), MBC 20080327, MG148437, MG148553, MG149012, MG148668, MG148783, MG148896, MG149127, MG149242, MG149357, MG149472; ***Zamia obliqua* A.Braun,** *Zamia obliqua* CO, Colombia (Chocó), Kiem & Norstog 36 (FTG), HS, MG148438, MG148554, MG149013, MG148669, MG148784, MG148897, MG149128, MG149243, MG149358, MG149473; ***Zamia obliqua* A.Braun,** *Zamia obliqua* PA1, Panama (Panamá), Calonje & Taylor PA09-020 (FTG), cult. Greg Holzman, MG148439, MG148555, MG149014, MG148670, MG148785, MG148898, MG149129, MG149244, MG149359, MG149474; ***Zamia obliqua* A.Braun,** *Zamia obliqua* PA2, Panama (Panamá), Bogler 1240 (FTG), FTBG 89162, MG148440, MG148556, MG149015, MG148671, MG148786, MG148899, MG149130, MG149245, MG149360, MG149475; ***Zamia oligodonta* E.Calderón & D.W.Stev.,** *Zamia oligodonta*, Colombia (Risaralda), Calderón-Sáenz 183 (NY), Cult. T. Maerowitz, MG148441, MG148557, MG149016, MG148672, MG148787, MG148900, MG149131, MG149246, MG149361, MG149476; ***Zamia onan-reyesii* C.Nelson & Sandoval,** *Zamia onan-reyesii*, Honduras (Cortés), Haynes et al. JLH03-045 (TEFH), MBC 20030887*V, MG148442, MG148558, MG149017, MG148673, MG148788, MG148901, MG149132, MG149247, MG149362, MG149477; ***Zamia paucijuga* Wieland,** *Zamia paucijuga* MX GRO, Mexico (Guerrero), Vovides et al. 1426 (XAL), MBC 20040444, MG148443, MG148562, MG149018, MG148677, MG148792, MG148905, MG149136, MG149251, MG149366, MG149481; ***Zamia paucijuga* Wieland,** *Zamia paucijuga* MX JAL, Mexico (Jalisco), Avendaño et al. 5255 (XAL), MBC 2000802*A, MG148444, MG148563, MG149020, MG148678, MG148793, MG148906, MG149137, MG149252, MG149367, MG149482; ***Zamia paucijuga* Wieland,** *Zamia paucijuga* MX OAX, Mexico (Oaxaca), Bogler 1198 (FTG), FTBG 931059, MG148446, MG148564, MG149021, MG148679, MG148794, MG148907, MG149138, MG149253, MG149368, MG149483; ***Zamia poeppigiana* Mart. & Eichler,** *Zamia poeppigiana*, Peru (Huanuco), Calonje D3 (FTG), cult. L. Whitelock, MG148447, MG148567, MG149022, MG148682, MG148797, MG148910, MG149141, MG149256, MG149371, MG149486; ***Zamia portoricensis* Urb.,** *Zamia portoricensis*, Puerto Rico, Meerow & Ayala-Silva 3105 (FTG), USDA-

ARS ZPR713, MG148448, MG148568, MG149023, MG148683, MG148798, MG148911, MG149142, MG149257, MG149372, MG149487; ***Zamia prasina* W.Bull,** *Zamia prasina* BZ, Belize (Cayo), Calonje et al. BZ08-009 (FTG), MBC 20080682, MG148449, MG148569, MG149024, MG148684, MG148799, MG148912, MG149143, MG149258, MG149373, MG149488; ***Zamia prasina* W.Bull,** *Zamia prasina* GU, Guatemala (Petén), Calonje FTBG18-01 (FTG), FTBG 941079, MG148450, MG148570, MG149025, MG148685, MG148800, MG148913, MG149144, MG149259, MG149374, MG149489; ***Zamia prasina* W.Bull,** *Zamia prasina* MX CHI, Mexico (Chiapas), Hubbuch & Walters 172C (FTG), FTBG 93931A, MG148451, MG148572, MG149026, MG148687, MG148802, MG148914, MG149146, MG149261, MG149376, MG149491; ***Zamia prasina* W.Bull,** *Zamia prasina* MX ROO, Mexico (Quintana Roo), Vovides et al. 1312 (XAL), MBC 2000799*A, MG148452, MG148573, MG149027, MG148688, MG148803, MG148916, MG149147, MG149262, MG149377, MG149492; ***Zamia prasina* W.Bull,** *Zamia prasina* MX YUC, Mexico (Yucatán), Vovides 856 (XAL), MBC 73385*B, MG148453, MG148574, MG149028, MG148689, MG148804, MG148917, MG149148, MG149263, MG149378, MG149493; ***Zamia pseudomonticola* L.D.Gómez ex D.W.Stev. & Sabato,** *Zamia pseudomonticola* CR, Costa Rica (Puntarenas), Calonje MAC04-012 (FTG), MBC 20041063, MG148454, MG148575, MG149029, MG148690, MG148805, MG148918, MG149149, MG149264, MG149379, MG149494; ***Zamia pseudomonticola* L.D.Gómez ex D.W.Stev. & Sabato,** *Zamia pseudomonticola* PA, Panama (Chiriquí), Espinosa 2011-004 (FTG), MBC 20020247*B, MG148455, MG148576, MG149031, MG148691, MG148806, MG148919, MG149150, MG149265, MG149380, MG149495; ***Zamia pseudoparasitica* J.Yates,** *Zamia pseudoparasitica* 1, Panama (Coclé), Stevenson DWS1206A (FTG, PMA, NY), MBC 2000319*A, MG148457, MG148577, MG149032, MG148692, MG148807, MG148920, MG149151, MG149266, MG149381, MG149496; ***Zamia pseudoparasitica* J.Yates,** *Zamia pseudoparasitica* 2, Panama (Veraguas), Kiem & Norstog 61 (FTG), FTBG 76663, MG148458, MG148578, MG149033, MG148693, MG148808, MG148921, MG149152, MG149267, MG149382, MG149497; ***Zamia pumila* L.,** *Zamia pumila* DR, Domican Republic, Calonje et al. DR09-001 (FTG), USDA-ARS ZDR111, MG148459, MG148579, MG149034, MG148694, MG148809, MG148922, MG149153, MG149268, MG149383, MG149498; ***Zamia purpurea* Vovides, J.D.Rees & Vázq.Torres,** *Zamia purpurea*, Mexico (Oaxaca),

Bogler 1197 (FTG), FTBG 93928, MG148460, MG148580, MG149035, MG148695, MG148810, MG148923, MG149154, MG149269, MG149384, MG149499; ***Zamia pygmaea* Sims,** *Zamia pygmaea*, Cuba (Isla de la Juventud), Calonje D2 (FTG), Cult. Loran Whitelock, MG148461, MG148581, MG149036, MG148696, MG148811, MG148924, MG149155, MG149270, MG149385, MG149500; ***Zamia pyrophylla* Calonje, D.W.Stev. & A.Lindstr.,** *Zamia pyrophylla*, Colombia (Chocó), Kress & Echeverry 89-2571 (US), MBC 20100027*A, MG148462, MG148469, MG149037, MG148584, MG148699, MG148925, MG149043, MG149158, MG149273, MG149388; ***Zamia restrepoi* (D.W.Stev.) A.Lindstr.,** *Zamia restrepoi*, Colombia (Cordoba), Calonje MBC18-03 (FTG), MBC 20100026*A, MG148463, MG148470, MG149038, MG148585, MG148700, MG148814, MG149044, MG149159, MG149274, MG149389; ***Zamia roezlii* Linden,** *Zamia roezlii* CO, Colombia (Valle del Cauca), Bogler 1224 (FTG), FTBG 94909, MG148464, MG148471, MG149039, MG148586, MG148701, MG148815, MG149045, MG149160, MG149275, MG149390; ***Zamia roezlii* Linden,** *Zamia roezlii* EC, Ecuador (Esmeraldas), Janovec et al. 1285 (NY), MBC 20001005*N, MG148465, MG148472, MG149040, MG148587, MG148702, MG148816, MG149046, MG149161, MG149276, MG149391; ***Zamia sandovalii* C.Nelson,** *Zamia sandovalii*, Honduras (Atlántida), Calonje MBC18-04 (FTG), MBC 95744*B, MG148466, MG148473, MG148928, MG148588, MG148703, MG148817, MG149047, MG149162, MG149277, MG149392; ***Zamia skinneri* Warsz. ex A.Dietr.,** *Zamia skinneri*, Panama (Bocas del Toro), Calonje MBC18-05 (FTG), MBC 20010794*A, MG148355, MG148474, MG148929, MG148589, MG148704, MG148818, MG149048, MG149163, MG149278, MG149393; ***Zamia soconuscensis* Schutzman, Vovides & Dehgan,** *Zamia soconuscensis*, Mexico (Chiapas), Bogler 1185 (FTG), FTBG 88313, MG148356, MG148475, MG148930, MG148590, MG148705, MG148819, MG149049, MG149164, MG149279, MG149394; ***Zamia sp.*,** *Zamia sp.* PE SAM, Peru (San Martin), S. W. Ingram 1953 (SEL), MS 1981-2029*B, MG148360, MG148477, MG148934, MG148592, MG148707, MG148821, MG149051, MG149166, MG149281, MG149396; ***Zamia spartea* A.DC.,** *Zamia spartea*, Mexico (Oaxaca), Walters et al. 9-3 (XAL), MBC 93927*J, MG148357, MG148476, MG148931, MG148591, MG148706, MG148820, MG149050, MG149165, MG149280, MG149395; ***Zamia splendens* Schutzman,** *Zamia splendens* 1, Mexico (Chiapas), Walters 2001-37-B (XAL), MBC 20010267*A,

MG148358, MG148524, MG148932, MG148639, MG148754, MG148867, MG149098, MG149213, MG149328, MG149443; ***Zamia splendens* Schutzman,** *Zamia splendens* 2, Mexico (Tabasco), Hubbuch & Walters 171 (FTG), FTBG 93930A, MG148359, MG148525, MG148933, MG148640, MG148755, MG148868, MG149099, MG149214, MG149329, MG149444; ***Zamia standleyi* Schutzman,** *Zamia standleyi*, Honduras (Cortés), Calonje FTBG18-02 (FTG), FTBG 65990C, MG148361, MG148478, MG148935, MG148593, MG148708, MG148822, MG149052, MG149167, MG149282, MG149397; ***Zamia stevensonii* A.S.Taylor & Holzman,** *Zamia stevensonii*, Panama (Panamá), Little & Stevenson 1168 (FTG), MBC 99885*A, MG148362, MG148480, MG148936, MG148595, MG148710, MG148823, MG149054, MG149169, MG149284, MG149399; ***Zamia tolimensis* Calonje, H.E.Esquivel & D.W.Stev.,** *Zamia tolimensis*, Colombia (Tolima), Little 8758 (US), HS, MG148363, MG148481, MG148937, MG148596, MG148711, MG148825, MG149055, MG149170, MG149285, MG149400; ***Zamia tuerckheimii* Donn.Sm.,** *Zamia tuerckheimii*, Guatemala (Alta Verapaz), Little & Stevenson 1163 (FTG), FTBG 771002B, MG148364, MG148482, MG148939, MG148597, MG148712, MG148826, MG149056, MG149171, MG149286, MG149401; ***Zamia urep* B.Walln.,** *Zamia urep*, Peru (Huánuco), Lindstrom s.n. (FTG, L, V), NN 15223A, MG148366, MG148483, MG148940, MG148598, MG148713, MG148827, MG149057, MG149172, MG149287, MG149402; ***Zamia variegata* Warsz.,** *Zamia variegata* BZ, Belize (Toledo), Calonje et al. BZ08-175 (FTG), MBC 20080675, MG148367, MG148565, MG148941, MG148680, MG148795, MG148908, MG149139, MG149254, MG149369, MG149484; ***Zamia variegata* Warsz.,** *Zamia variegata* GU, Guatemala (Izabal), Standley 23873 (US), FTBG 73196, MG148368, MG148566, MG148942, MG148681, MG148796, MG148909, MG149140, MG149255, MG149370, MG149485; ***Zamia vazquezii* D.W.Stev., Sabato & De Luca,** *Zamia vazquezii* MX VER, Mexico (Veracruz), Gonzalo Castillo CC4480 (XAL), JBC 1986-003, MG148369, MG148484, MG148943, MG148599, MG148714, MG148828, MG149058, MG149173, MG149288, MG149403; ***Zamia wallisii* Braun,** *Zamia wallisii*, Colombia (Antioquia), Little & Stevenson 1165 (FTG), MBC 20010301*A, MG148370, MG148485, MG148944, MG148600, MG148715, MG148829, MG149059, MG149174, MG149289, MG149404.

APPENDIX B: MAXIMUM PARSIMONY AND MAXIMUM LIKELIHOOD TREES

Fig. B1. Best scoring tree from maximum likelihood analysis conducted with RAxML. Nodes annotated with bootstrap values.

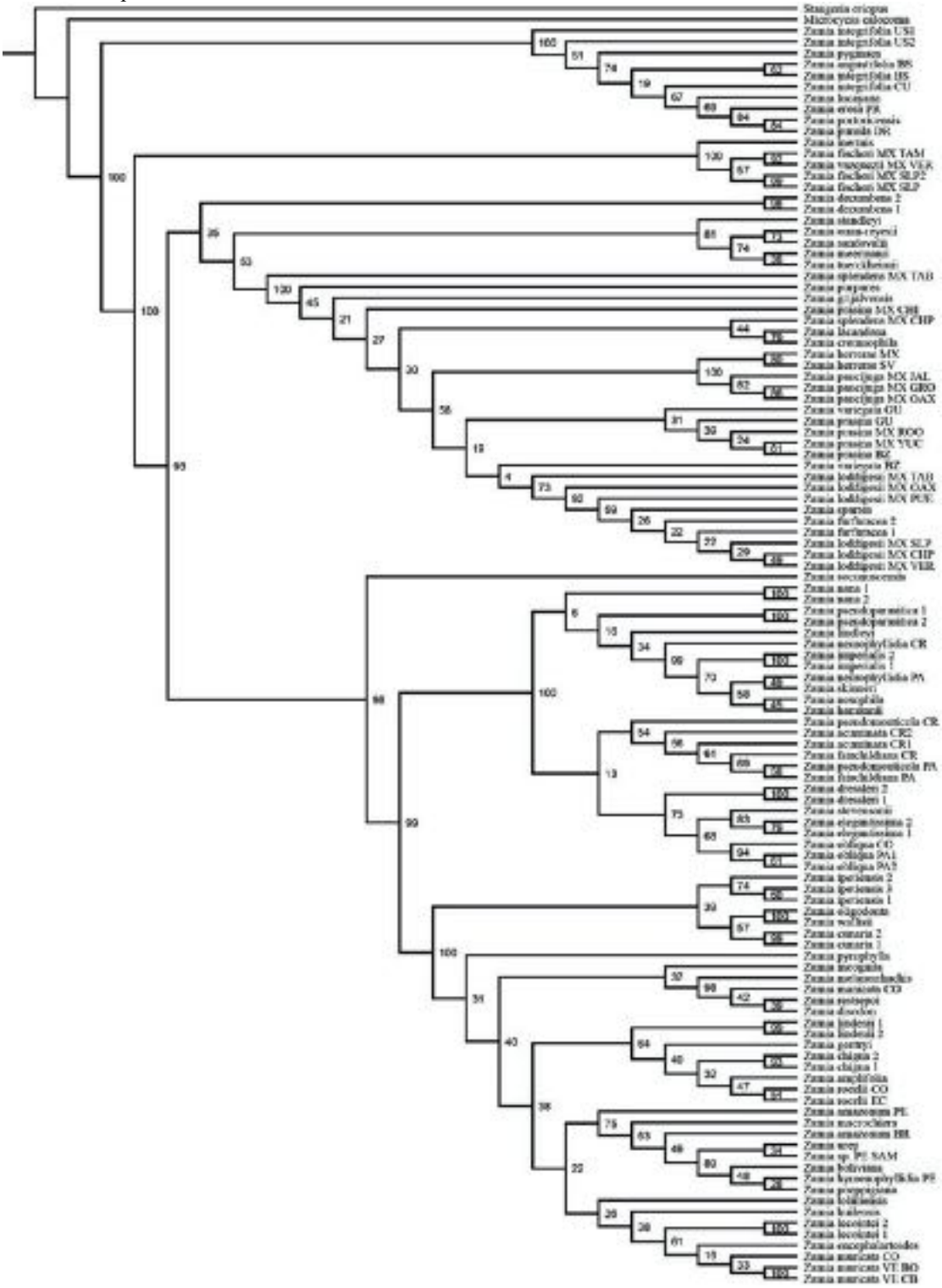


Fig. B2. 50% majority rule consensus tree of 1820 equally parsimonious trees found by heuristic parsimony analysis of a concatenated matrix of sequences from nine single-copy nuclear gene loci and one chloroplast gene locus. Nodes annotated with parsimony jacknife support percentages.

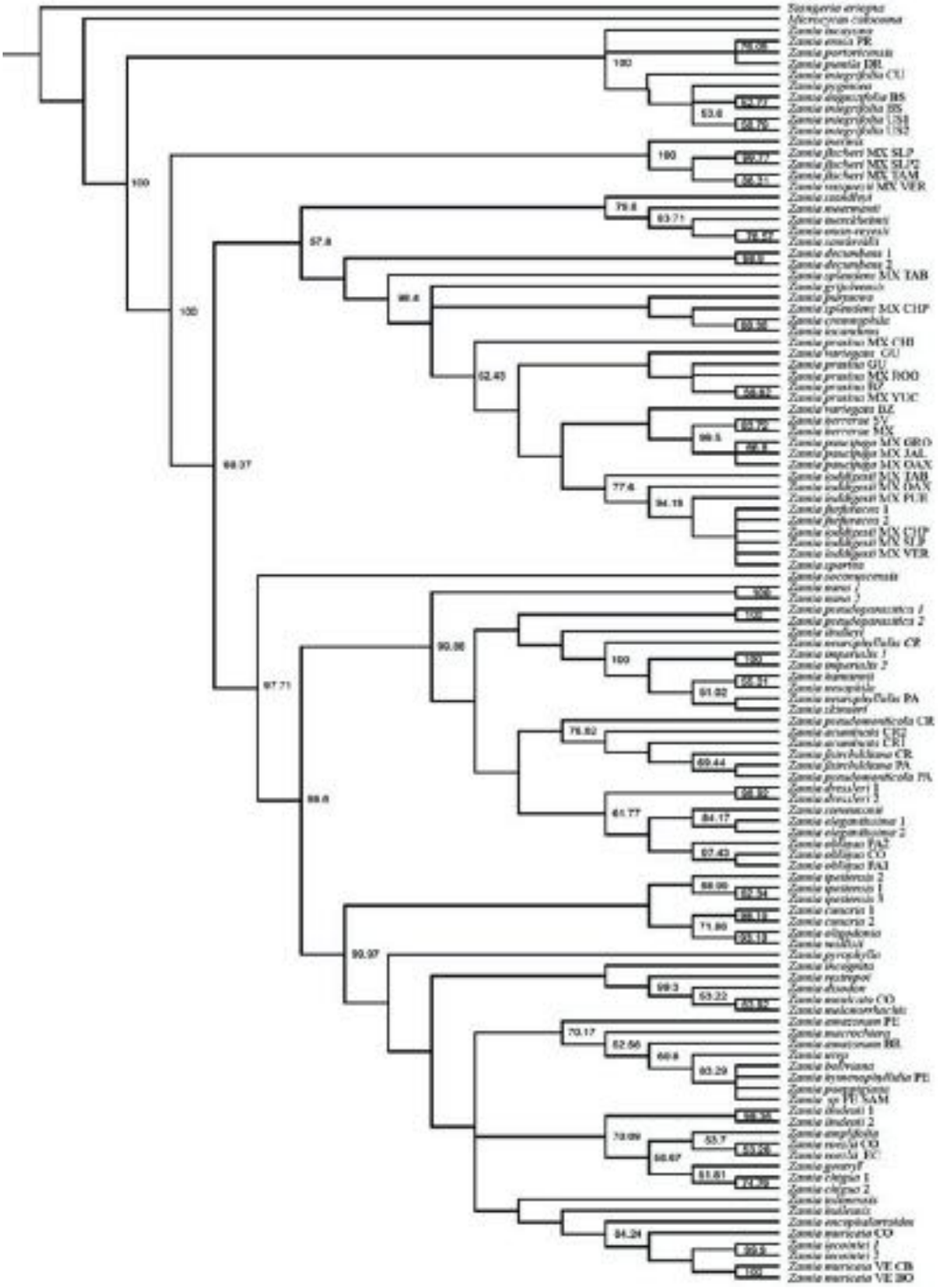


Fig. B3. Maximum parsimony strict consensus tree for locus *40s*. Branches annotated with parsimony jacknife support percentages.

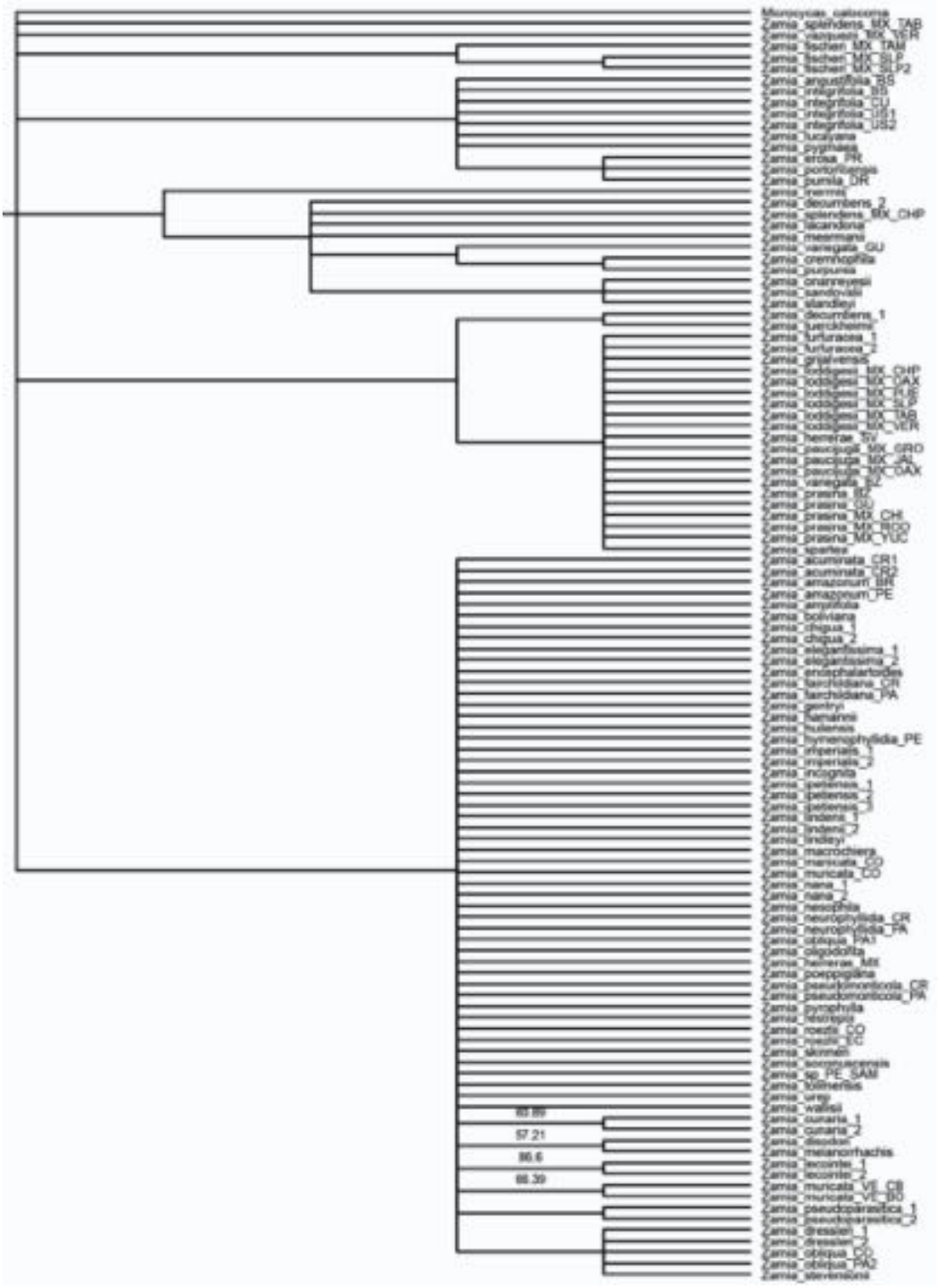


Fig. B4. Maximum parsimony strict consensus tree for locus *ATG2*. Branches annotated with parsimony jacknife support percentages.

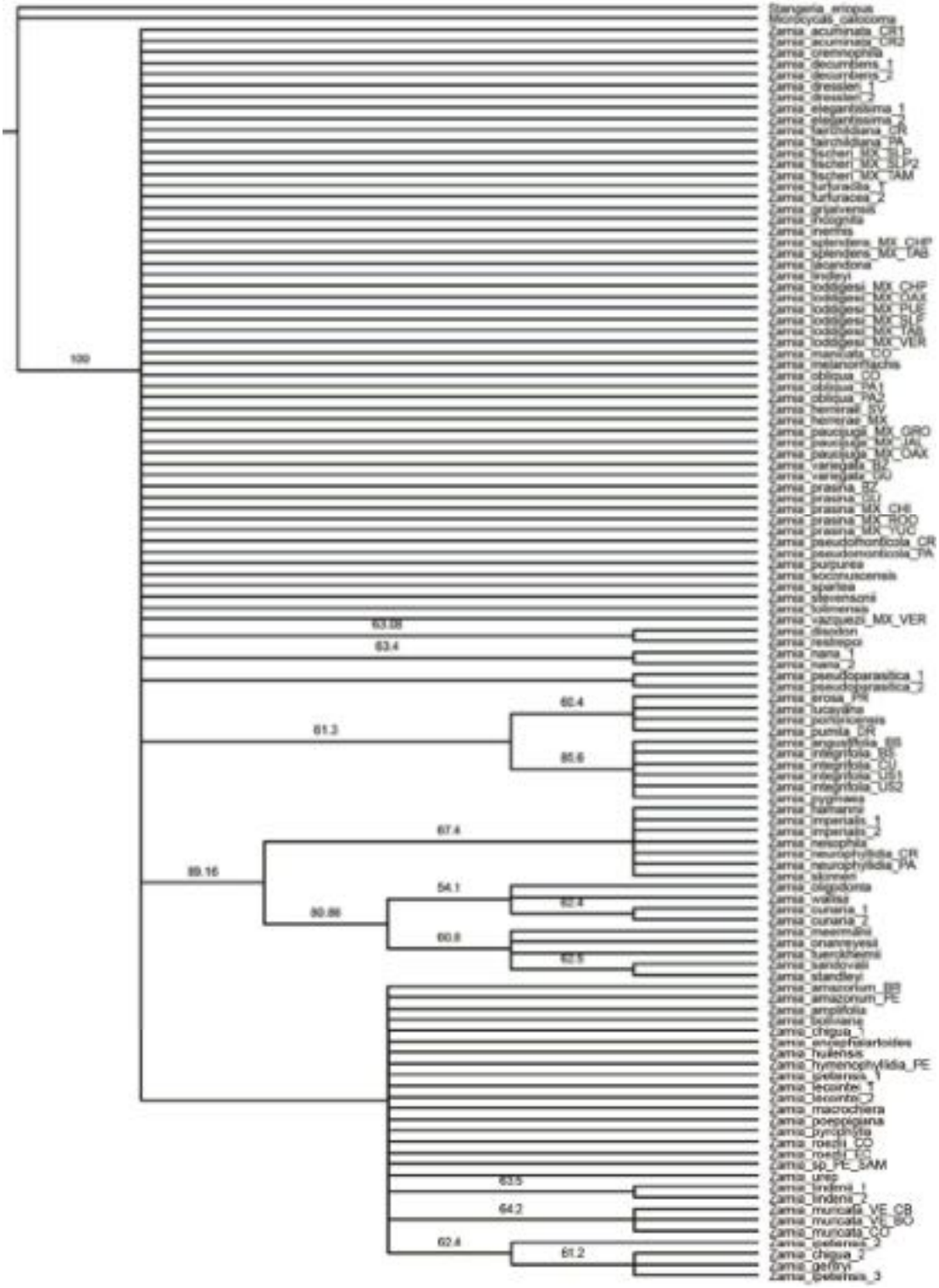


Fig. B5. Maximum parsimony strict consensus tree for locus *CyAG*. Branches annotated with parsimony jacknife support percentages.

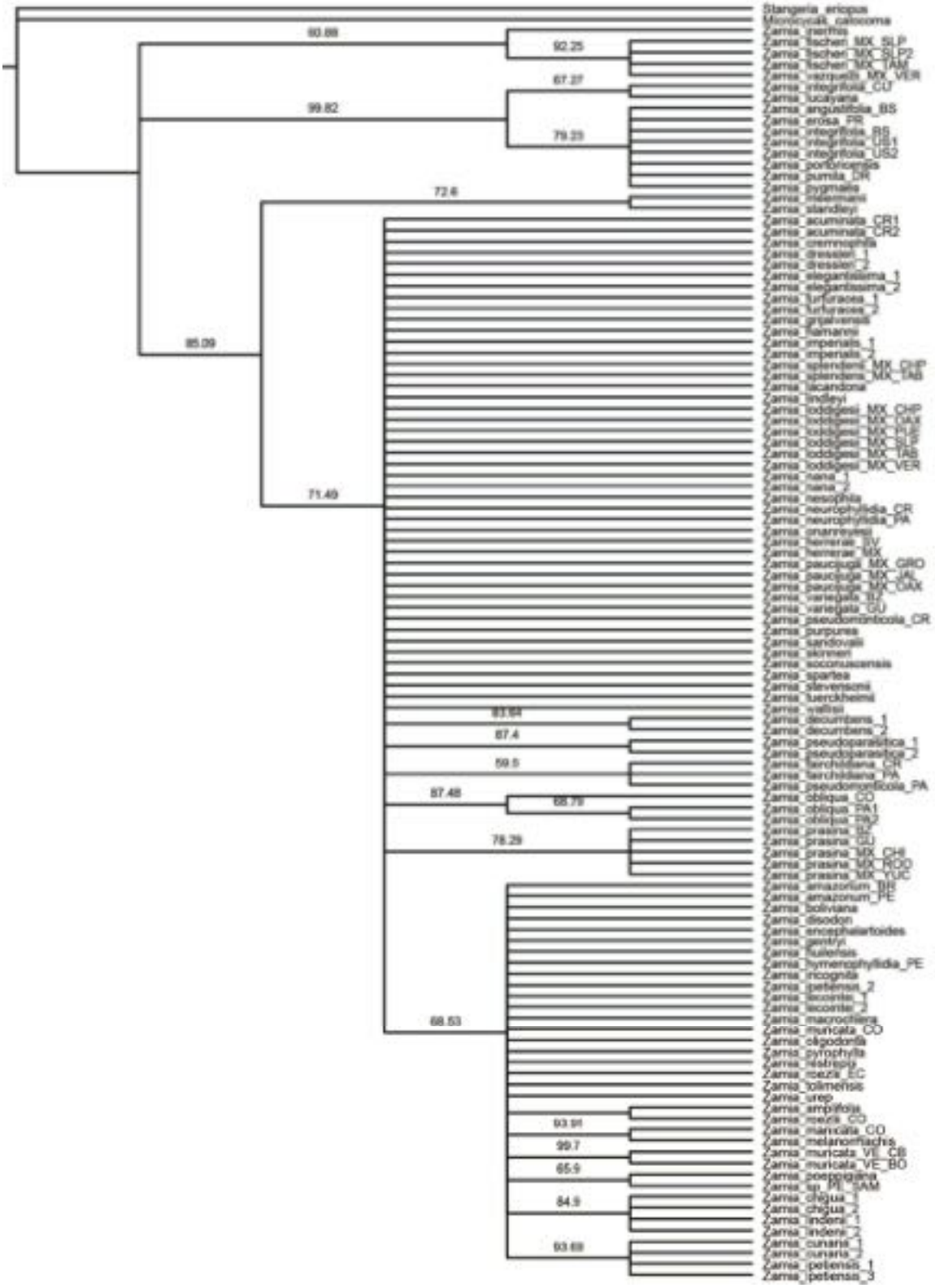


Fig. B6. Maximum parsimony strict consensus tree for locus GroES. Branches annotated with parsimony jacknife support percentages.

Fig. B7. Maximum parsimony strict consensus tree for locus HTS. Branches annotated with parsimony jacknife support percentages.

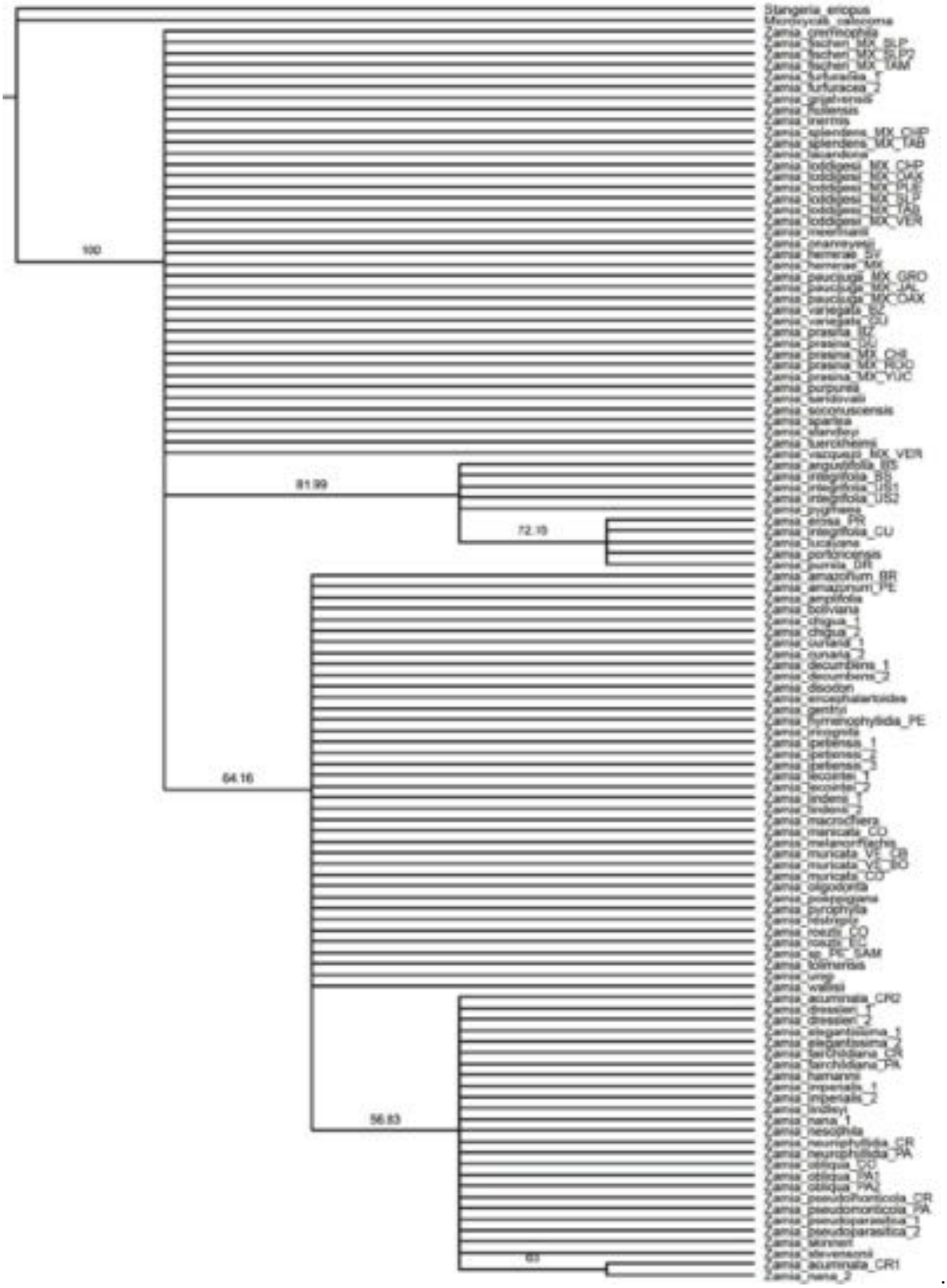


Fig. B8. Maximum parsimony strict consensus tree for locus LiSH. Branches annotated with parsimony jacknife support percentages.

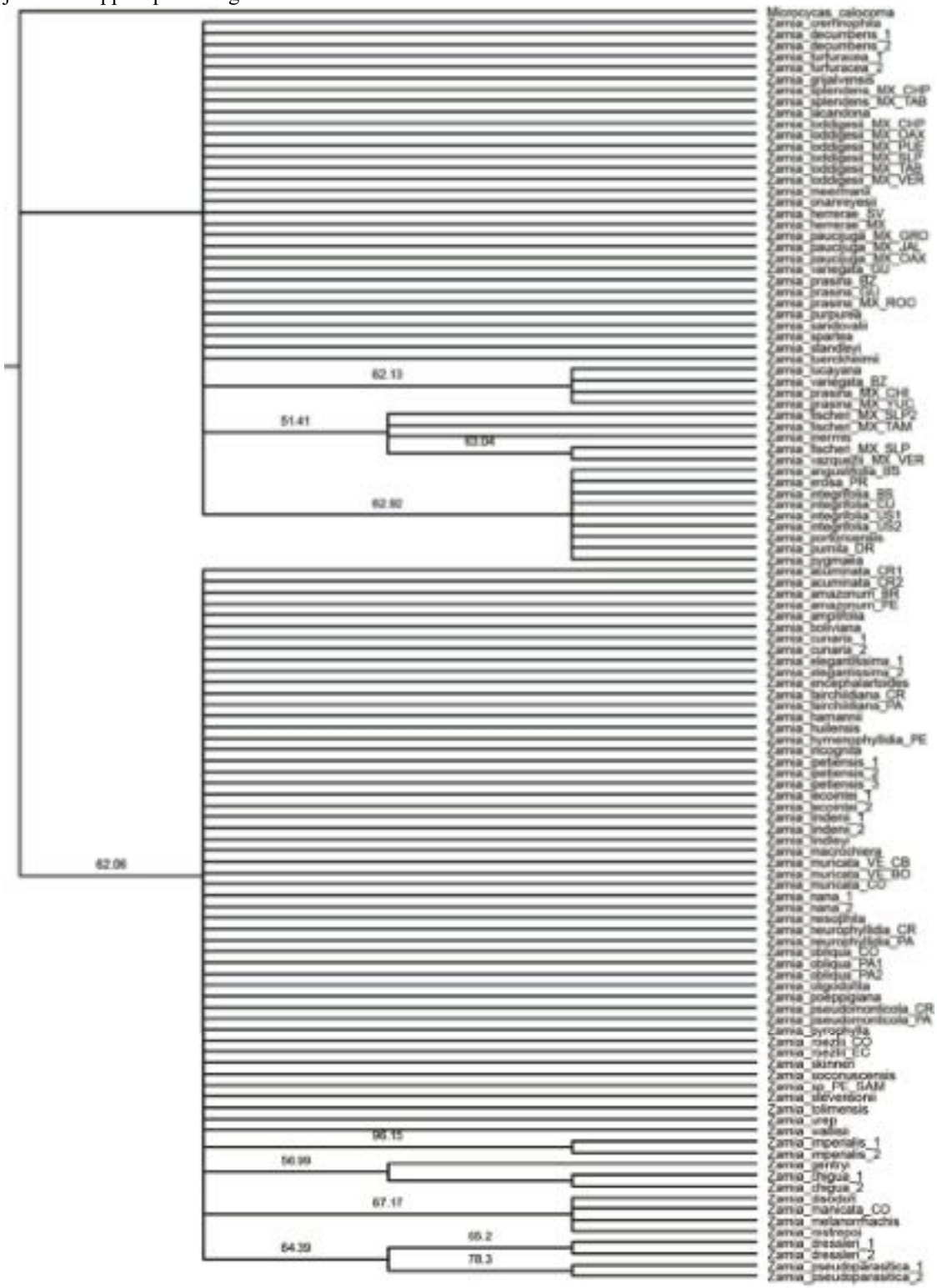


Fig. B9. Maximum parsimony strict consensus tree for locus PEX4. Branches annotated with parsimony jacknife support percentages.

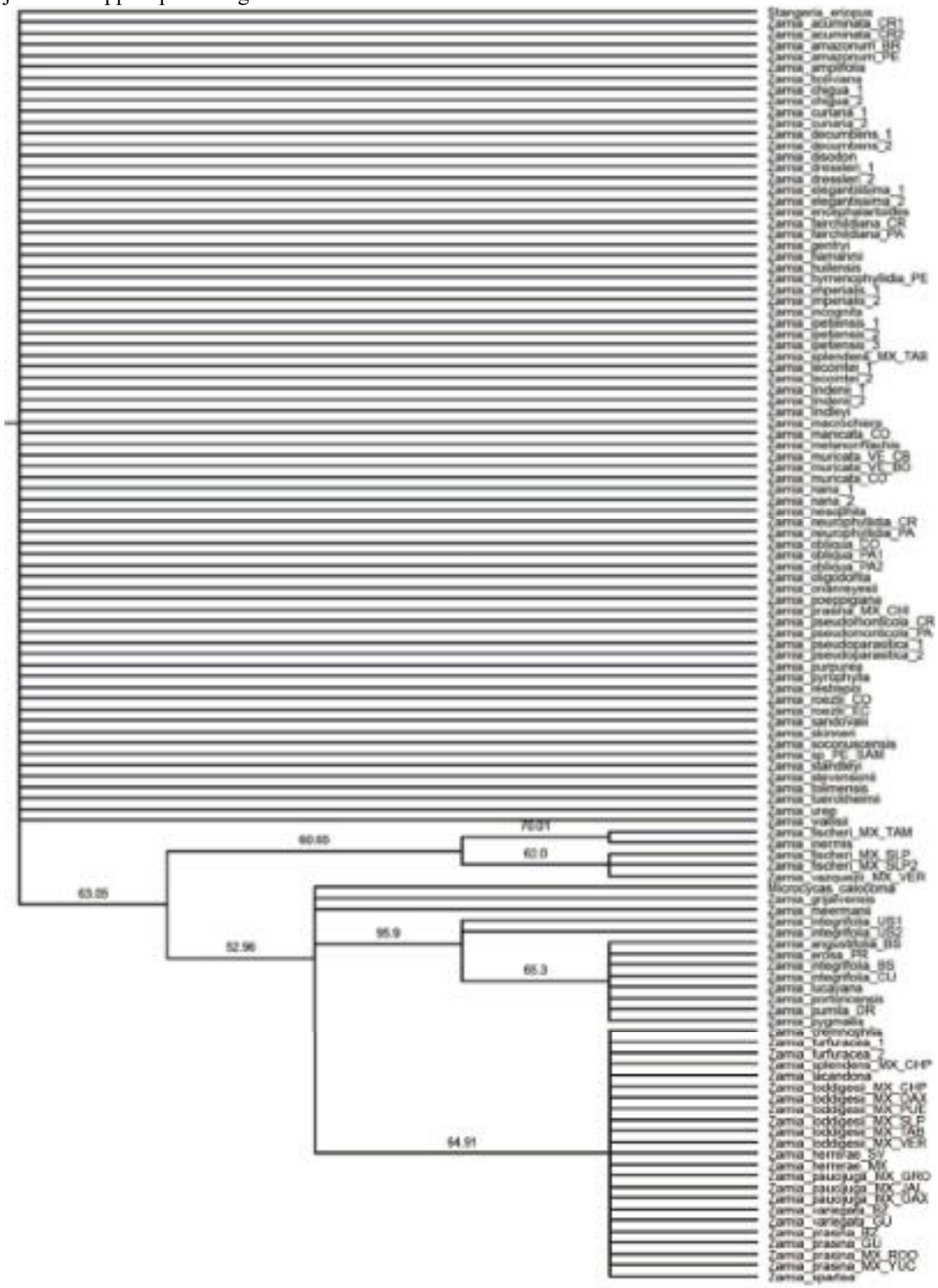


Fig. B10. Maximum parsimony strict consensus tree for locus PMP22. Branches annotated with parsimony jacknife support percentages.

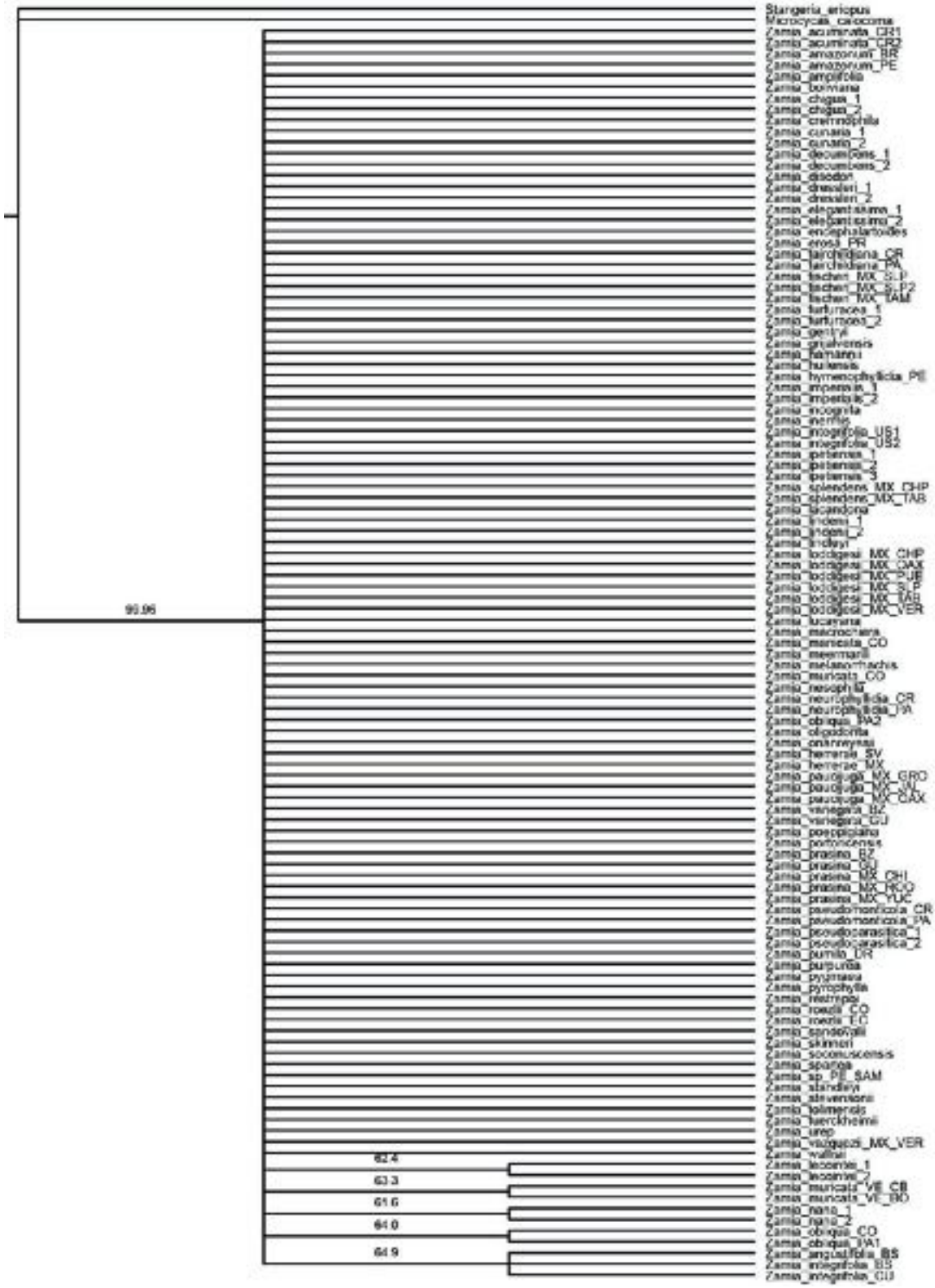


Fig. B11. Maximum parsimony strict consensus tree for locus PsbK/I. Branches annotated with parsimony jacknife support percentages.

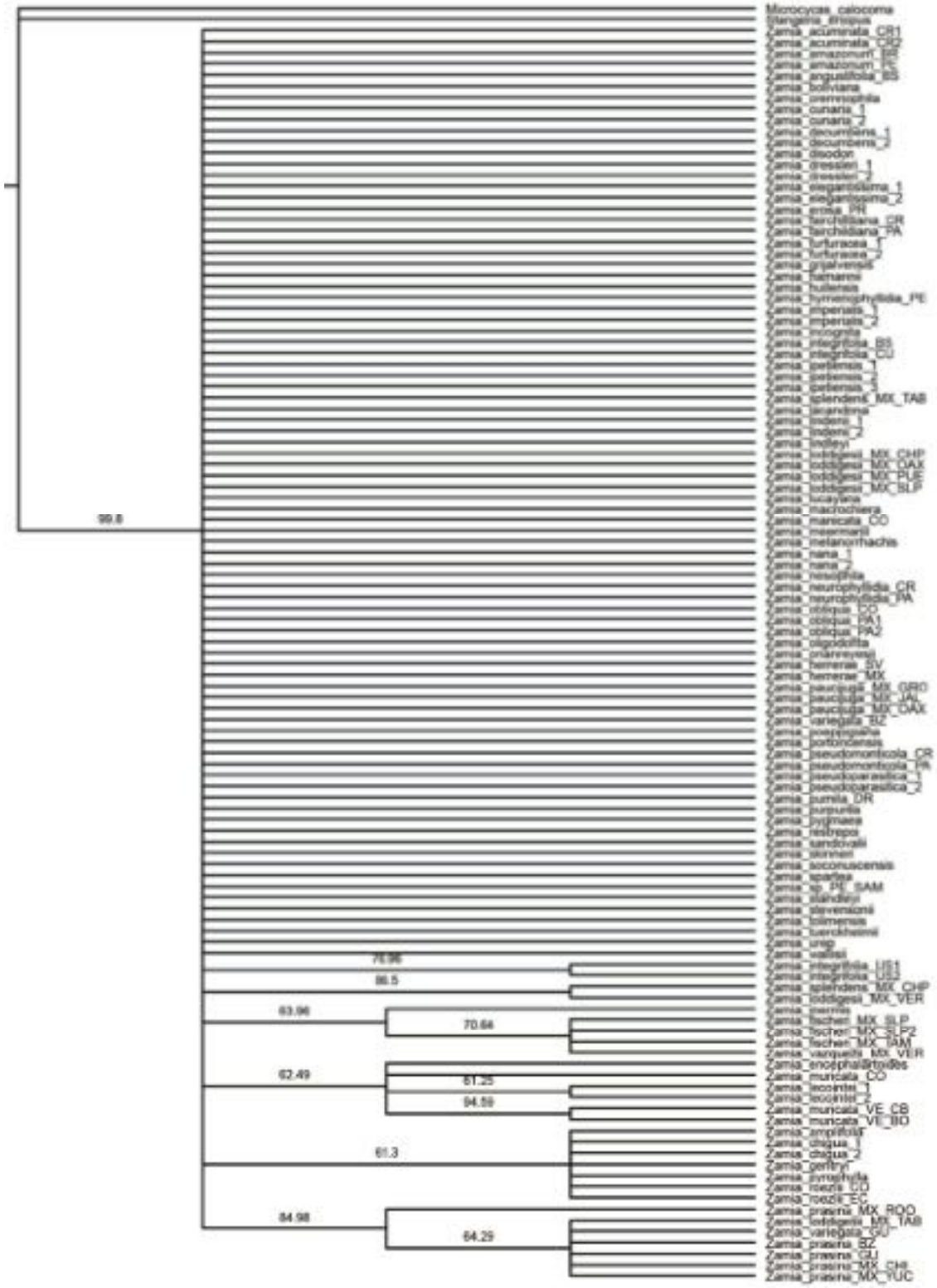
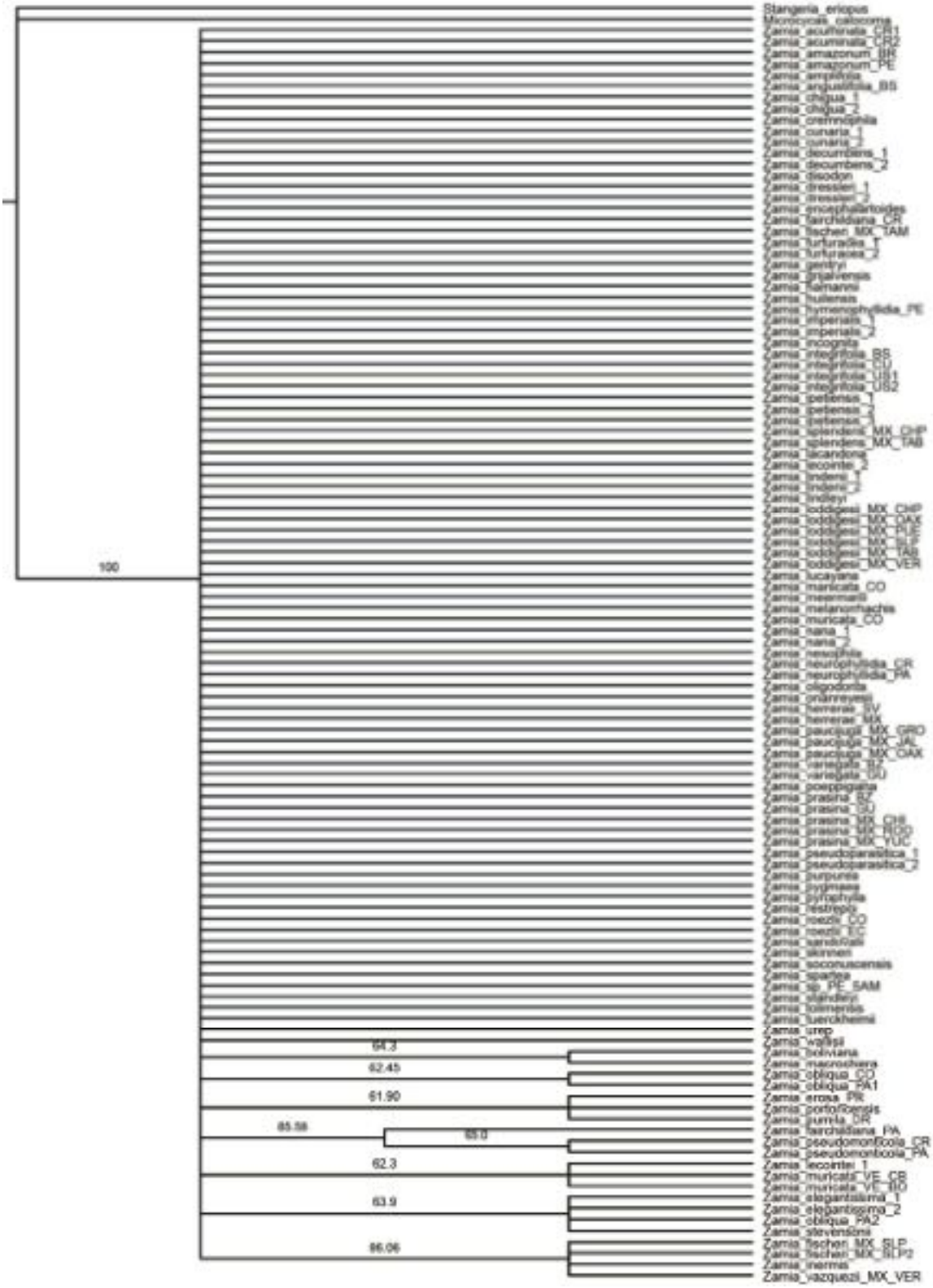


Fig. B12. Maximum parsimony strict consensus tree for locus WRKY4. Branches annotated with parsimony jacknife support percentages.



APPENDIX C: MAXIMUM CLADE CREDIBILITY (MCC) CONSENSUS GENE TREES FOR TEN INDIVIDUAL LOCI DERIVED FROM *BEAST SPECIES TREE ANALYSIS. INDIVIDUAL TREES WERE CREATED BY COMBINING SINGLE-LOCUS OUTPUTS FOR TWO INDEPENDENT *BEAST RUNS OF 1.1 BILLION MCMC ITERATIONS EACH WITH TREES SAMPLED EVERY 5,000 ITERATIONS, RESULTING IN A TOTAL OF 440,000 TREES USED IN EACH MCC GENE TREE ESTIMATION.

Fig. C1. Maximum clade credibility tree for locus *40S*. Numbers at nodes are posterior probability scores.

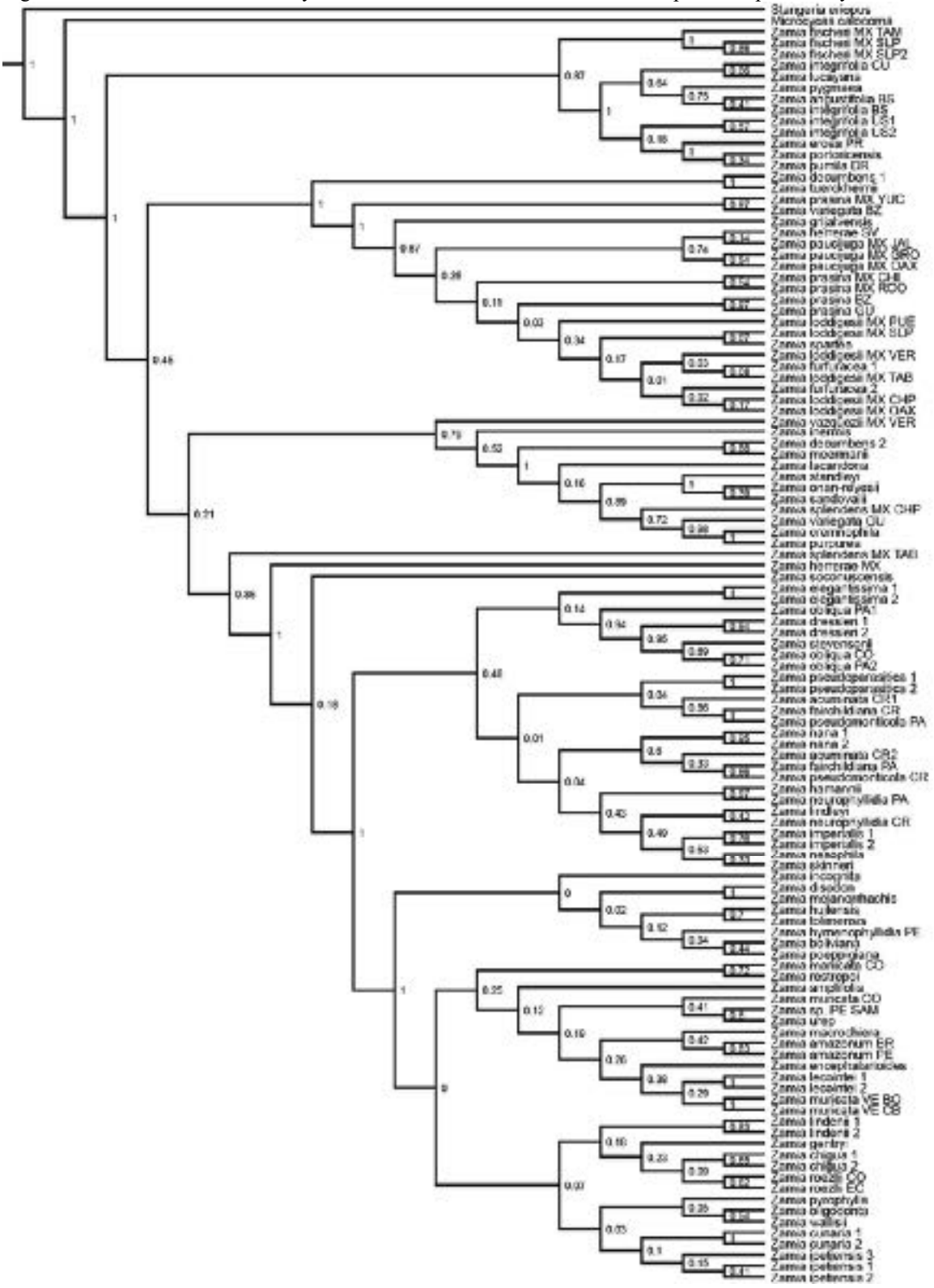


Fig. C2. Maximum clade credibility tree for locus *ATG2*. Numbers at nodes are posterior probability scores.

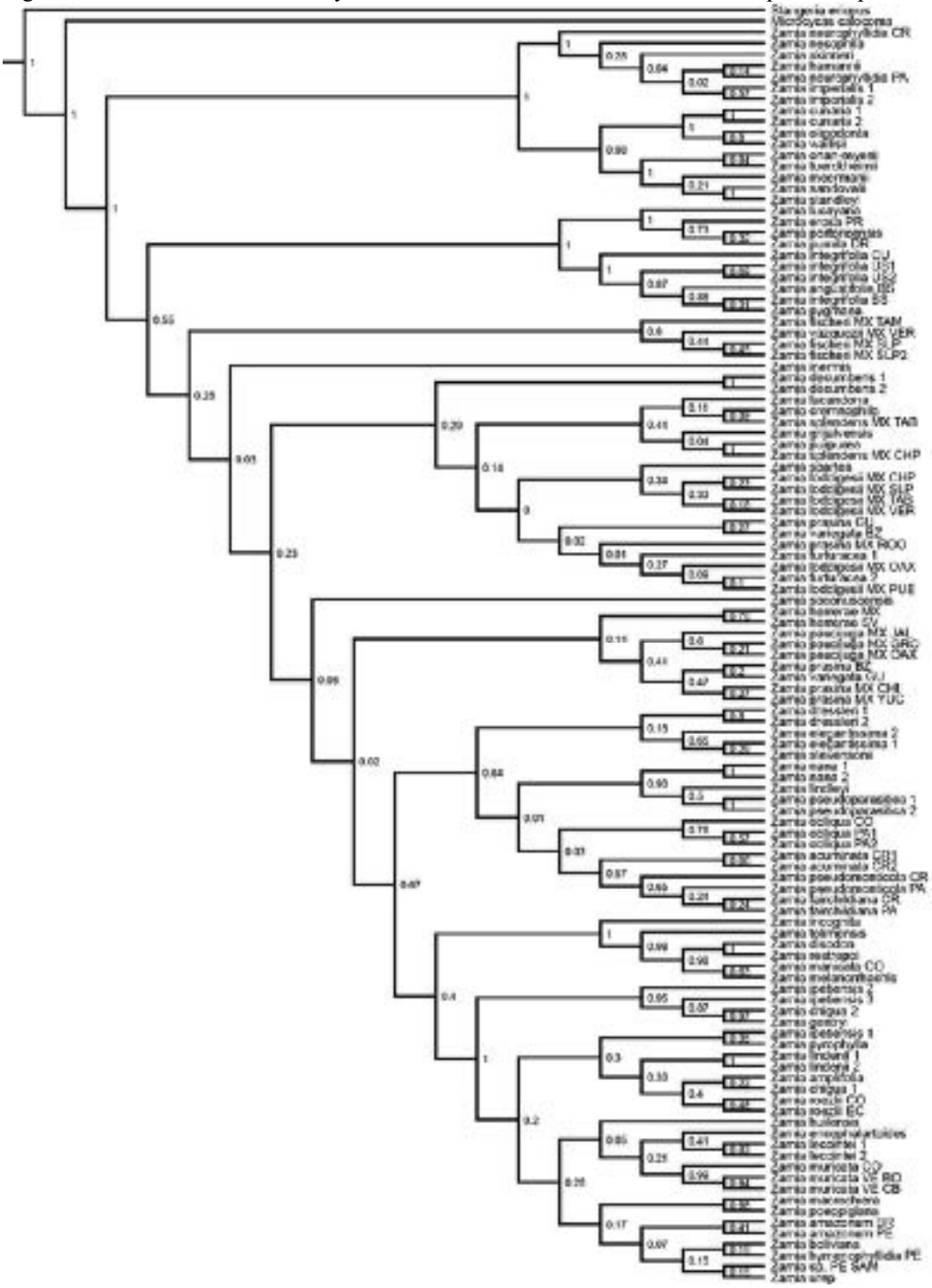


Fig. C3. Maximum clade credibility tree for locus *CyAG*. Numbers at nodes are posterior probability scores.

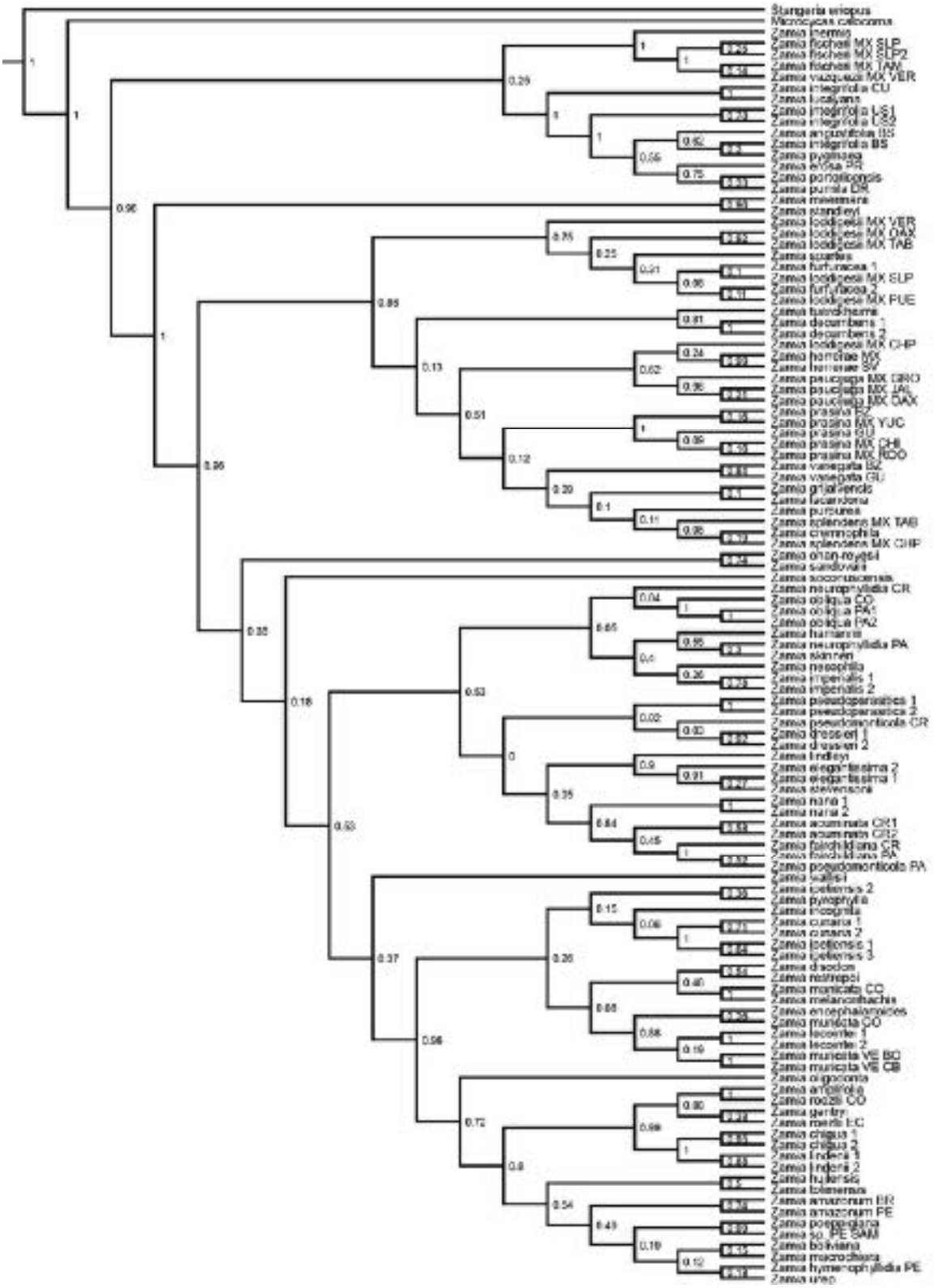


Fig. C4. Maximum clade credibility tree for locus *GroES*. Numbers at nodes are posterior probability scores.

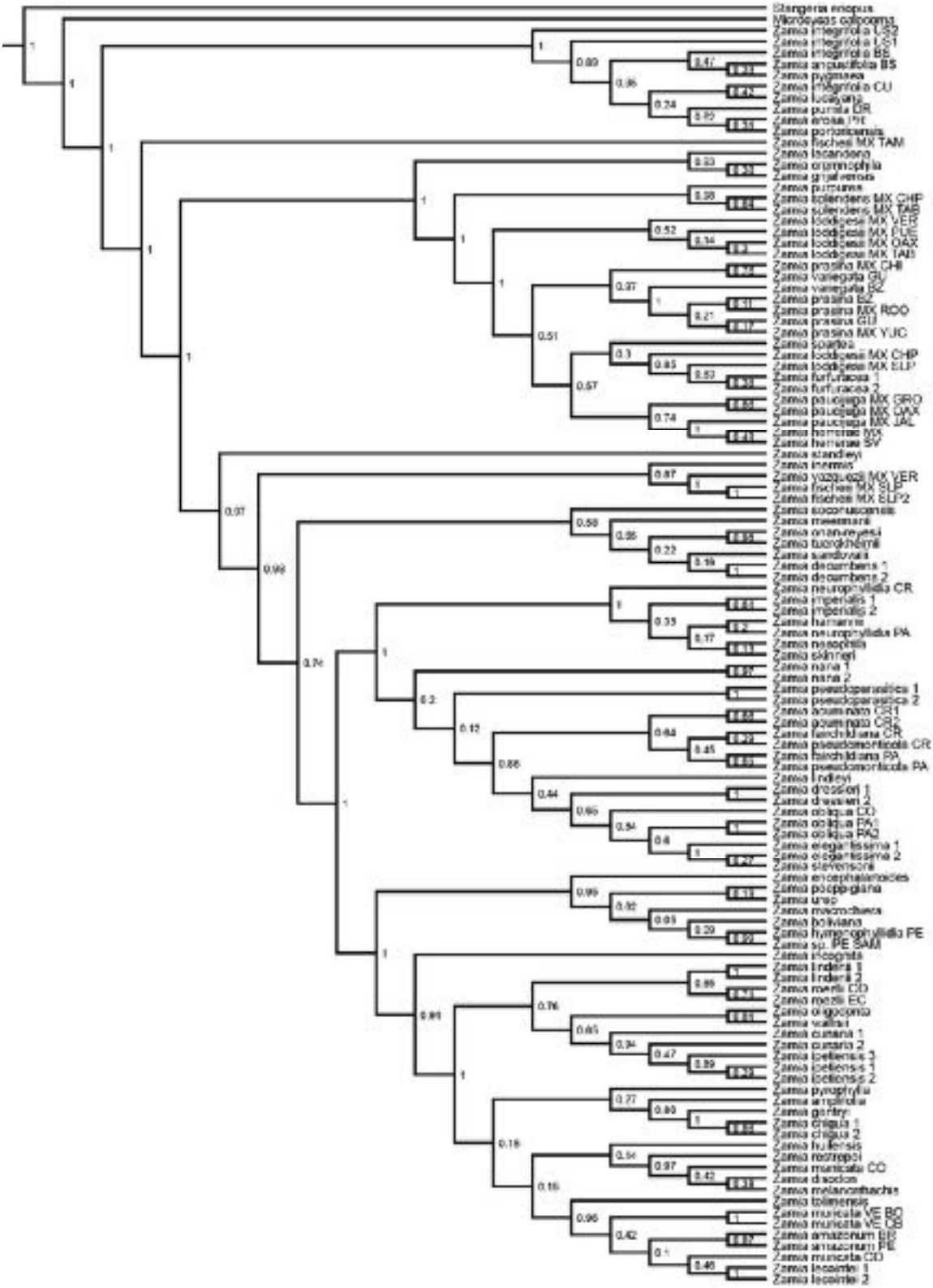


Fig. C5. Maximum clade credibility tree for locus *HTS*. Numbers at nodes are posterior probability scores.

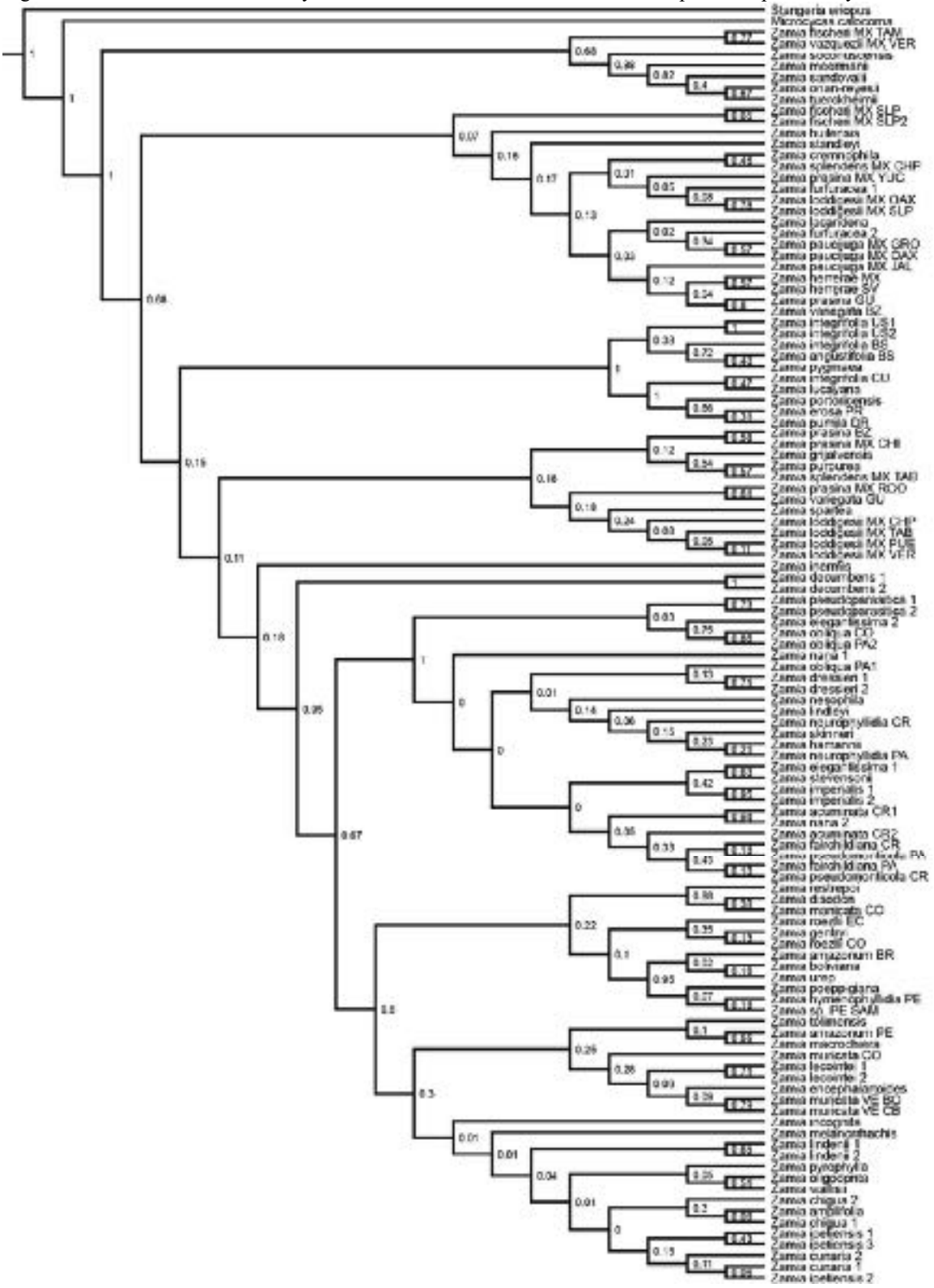


Fig. C6**.** Maximum clade credibility tree for locus *LiSH*. Numbers at nodes are posterior probability scores.

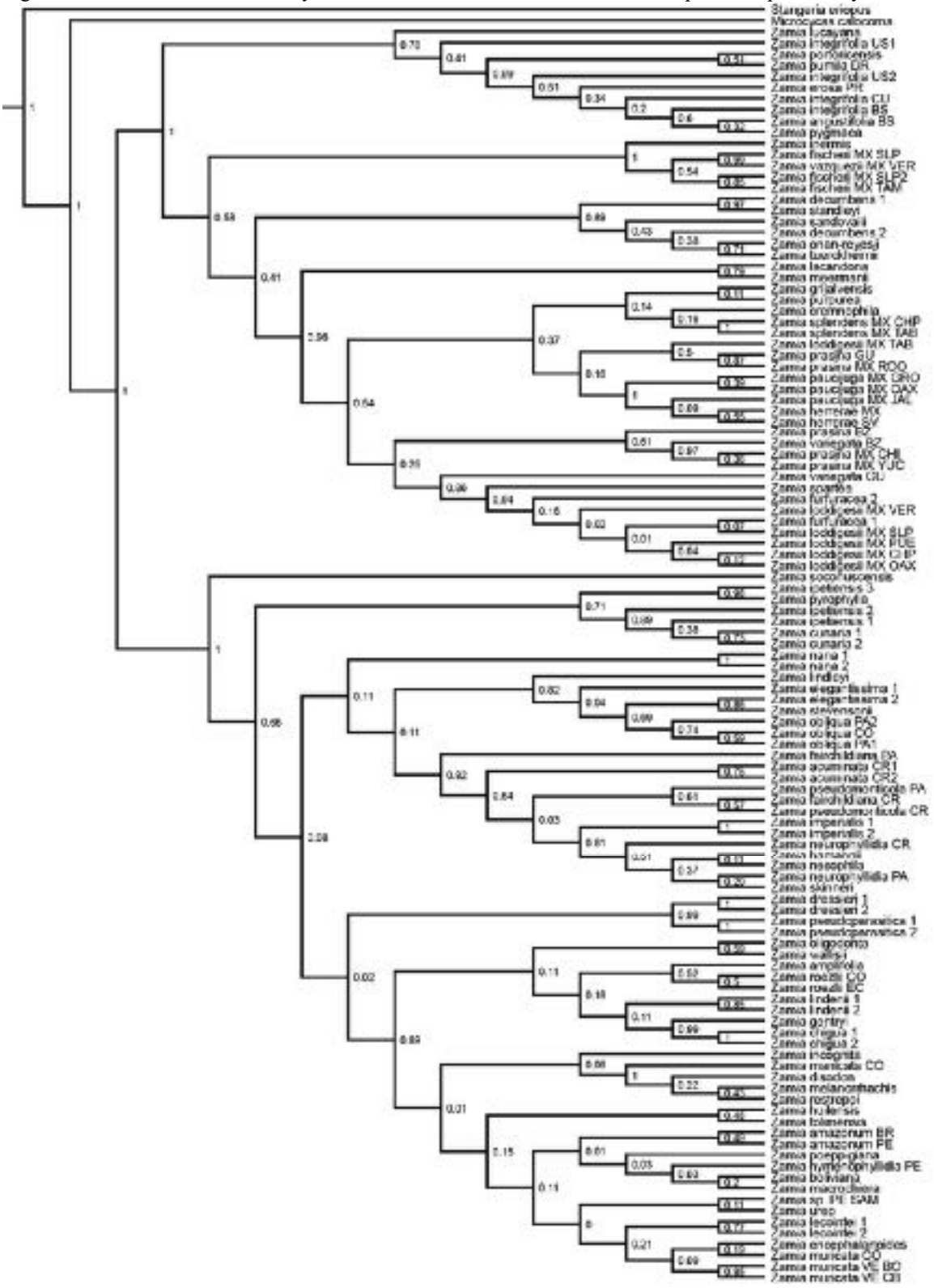


Fig. C7. Maximum clade credibility tree for locus *PEX4*. Numbers at nodes are posterior probability scores.

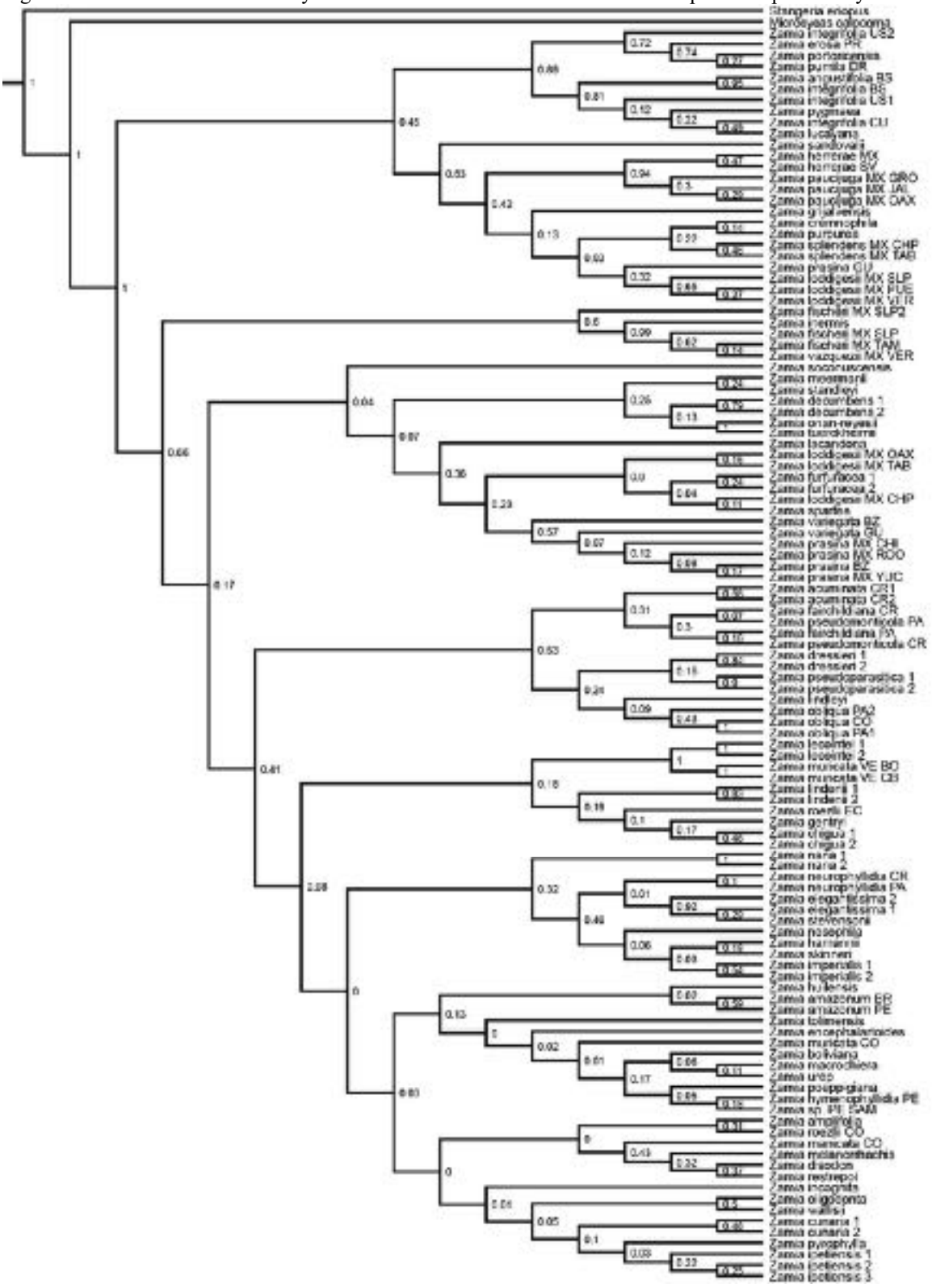


Fig. C8**.** Maximum clade credibility tree for locus *PMP22*. Numbers at nodes are posterior probability scores.

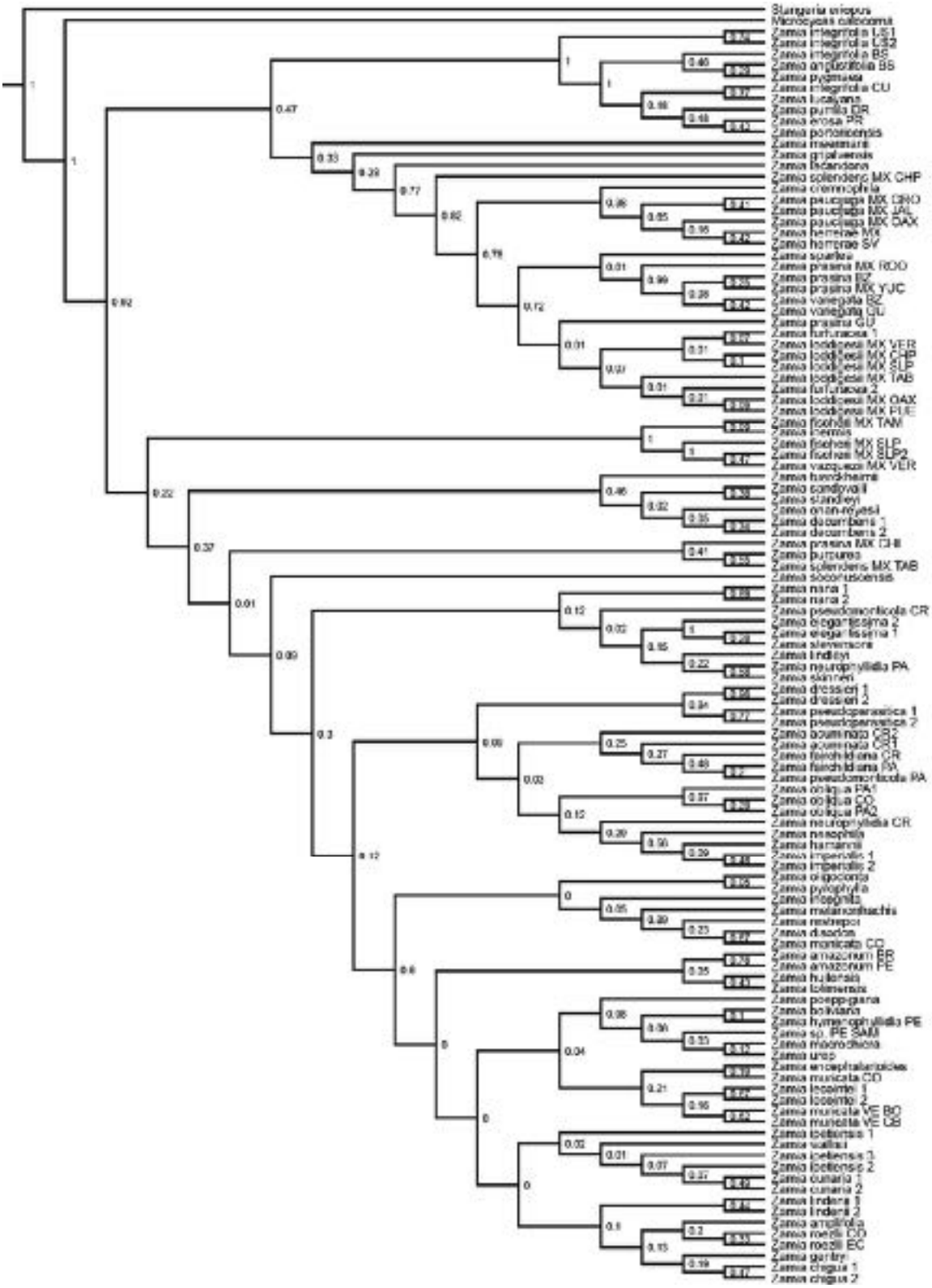


Fig. C9. Maximum clade credibility tree for locus *PsbK/I*. Numbers at nodes are posterior probability scores.

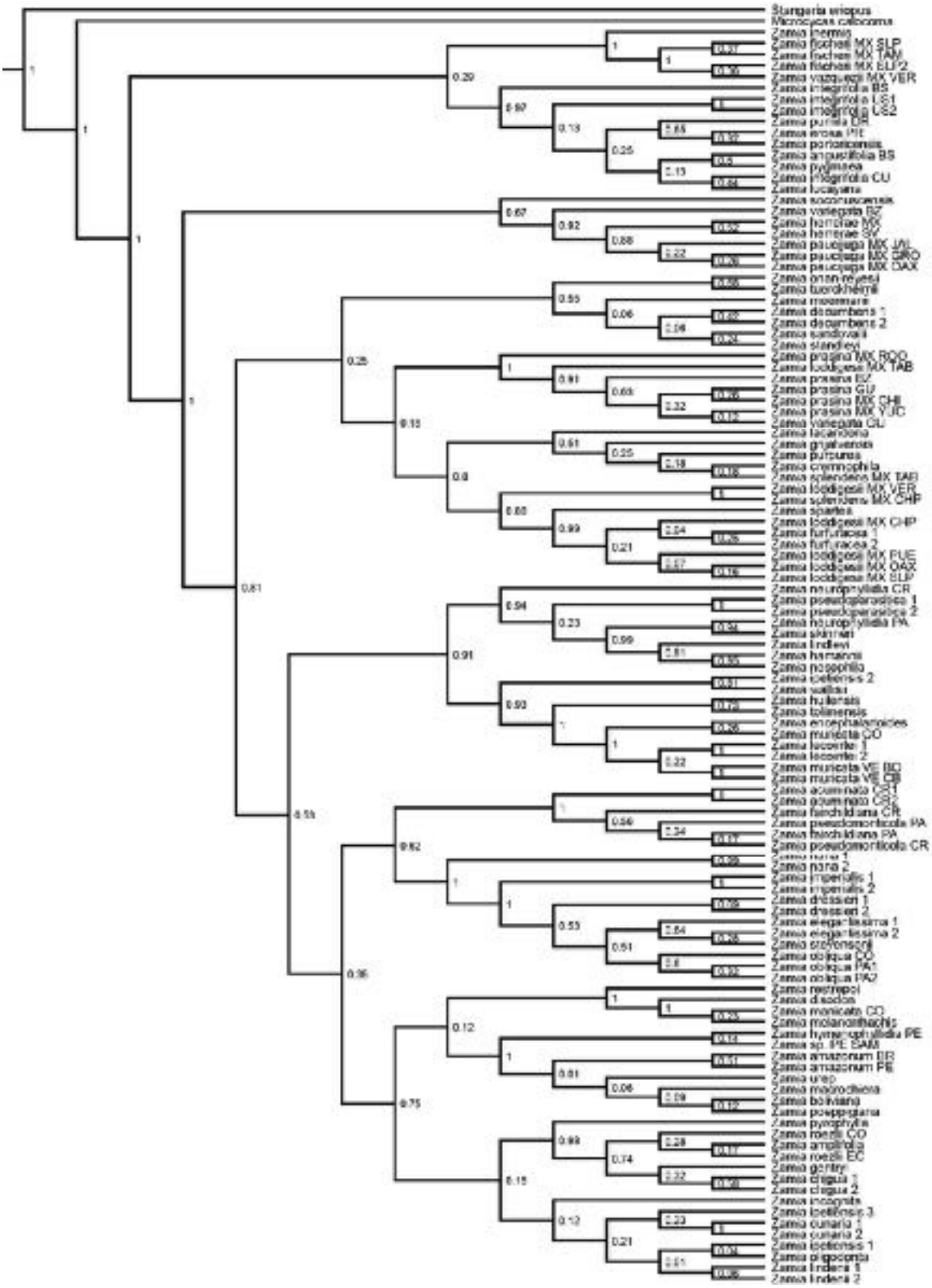
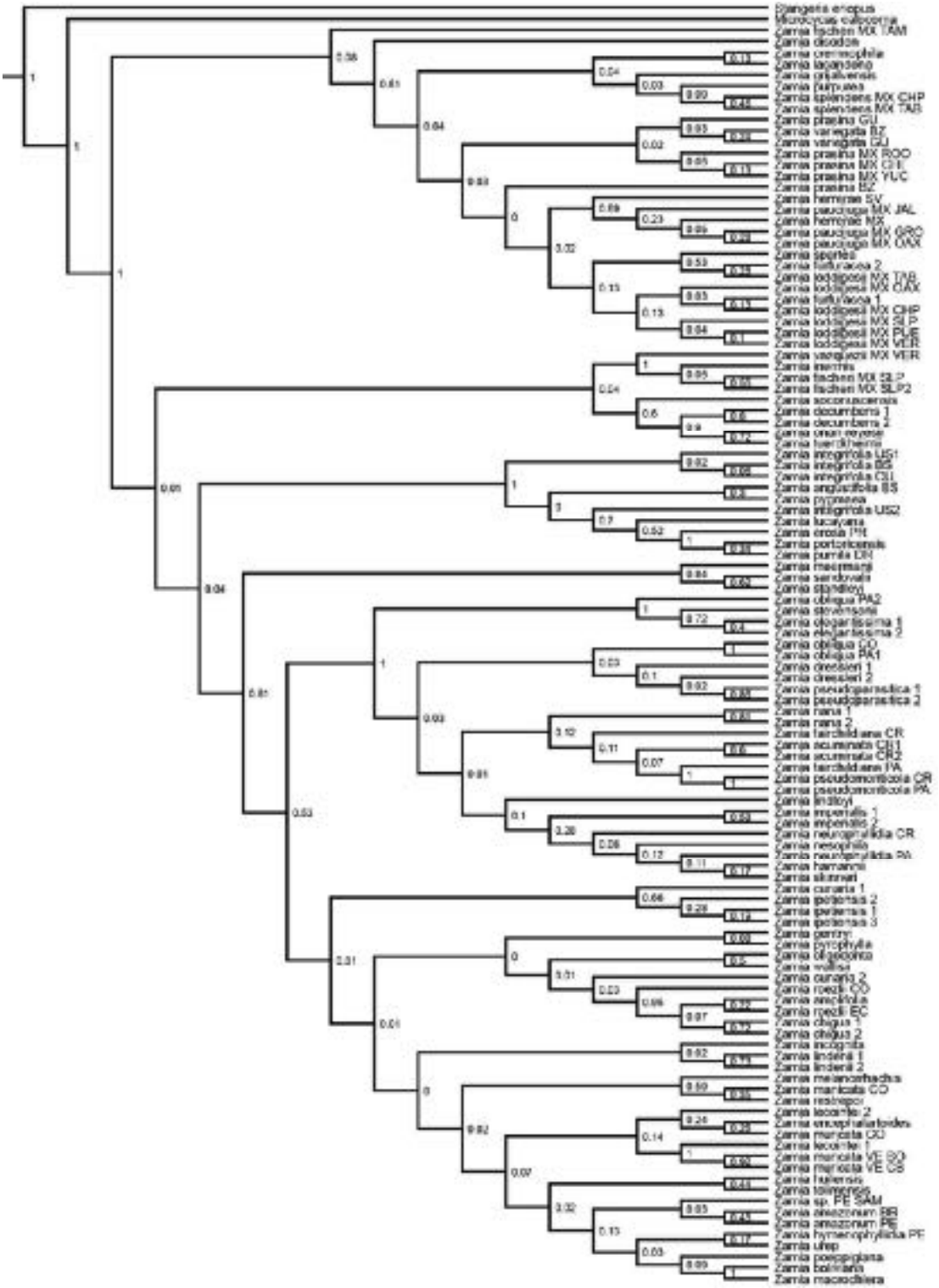


Fig. C10**.** Maximum clade credibility tree for locus *WRKY4*. Numbers at nodes are posterior probability scores.



APPENDIX D: DIVERSIFICATION ANALYSES FIGURES

Fig. D1. Lineage Through Time (LTT) plots for the stem and crown nodes of *Zamia* and major clades. Note differences in timescales and lineage numbers

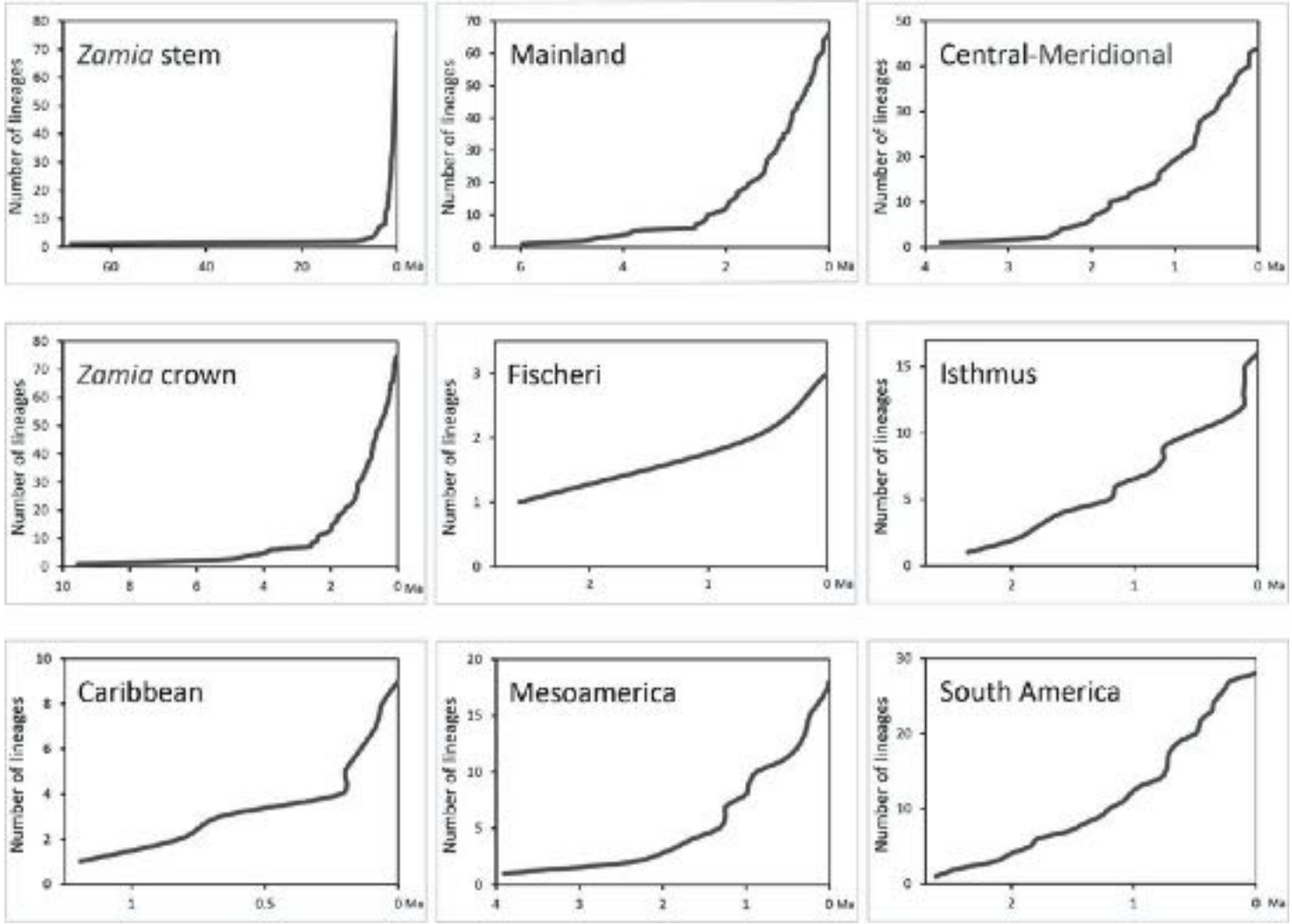
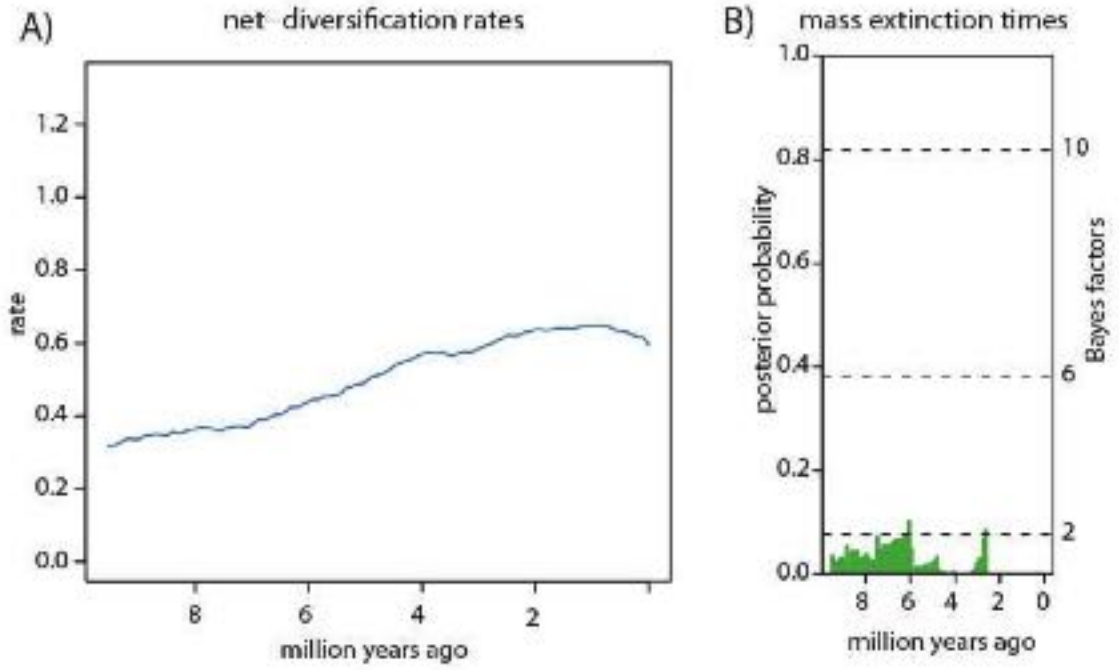


Fig. D2. Results from the Comet analysis under the TESS model. A) Net-diversification rates though time for the crown group of *Zamia*, median shown. The increase in diversification through time does not correspond to significant shifts in speciation or extinction rates. B) Results from the mass extinction event model. Two marginally supported mass extinctions are inferred to happen at around 6 mya and 2.5 mya



APPENDIX E: ANCESTRAL CHARACTER RECONSTRUCTIONS

Fig. E1. Ancestral character state reconstruction for arborescence

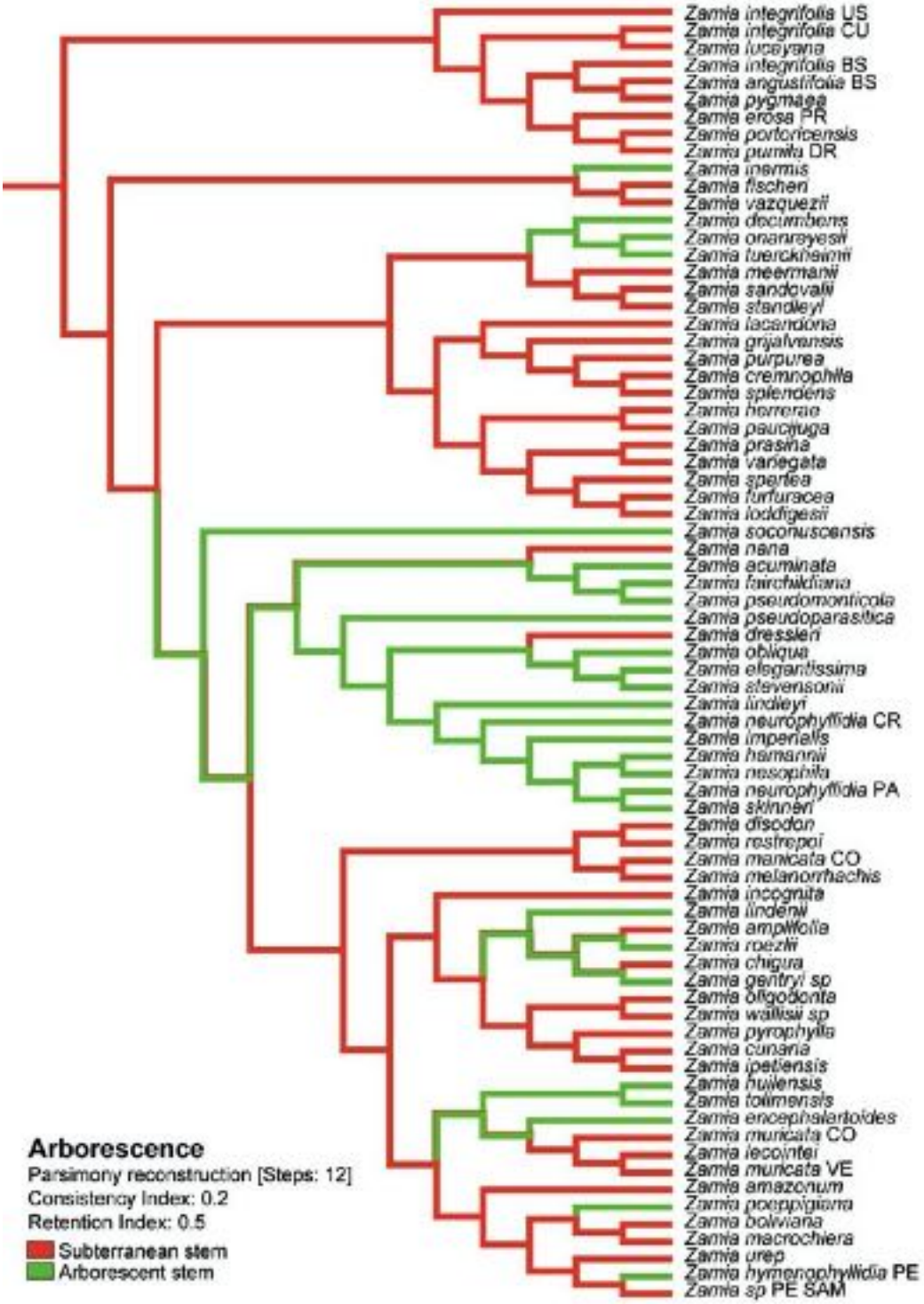


Fig. E2. Ancestral character state reconstruction for petiole prickle presence.

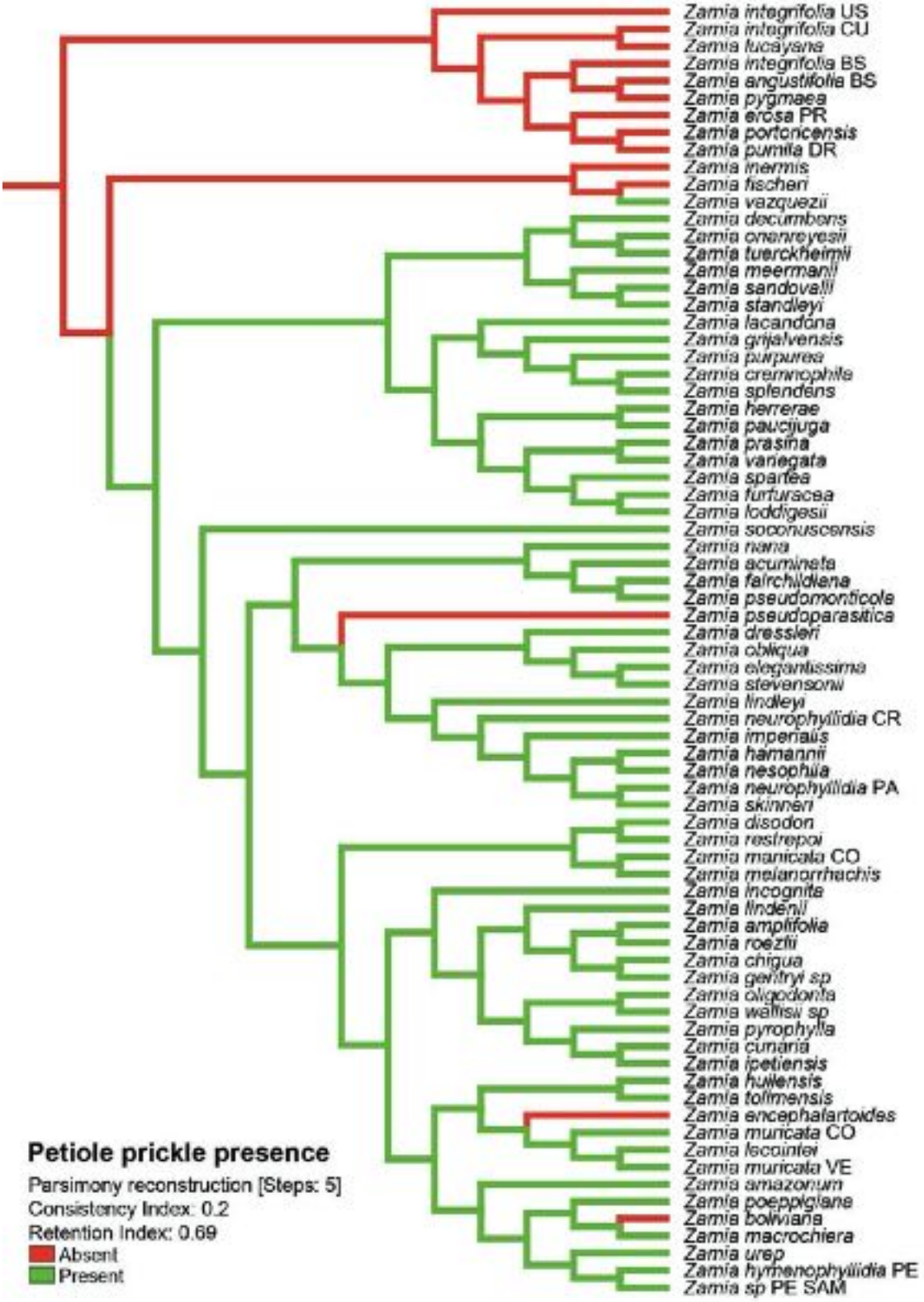


Fig. E3. Ancestral character state reconstruction for leaflet teeth prominence.

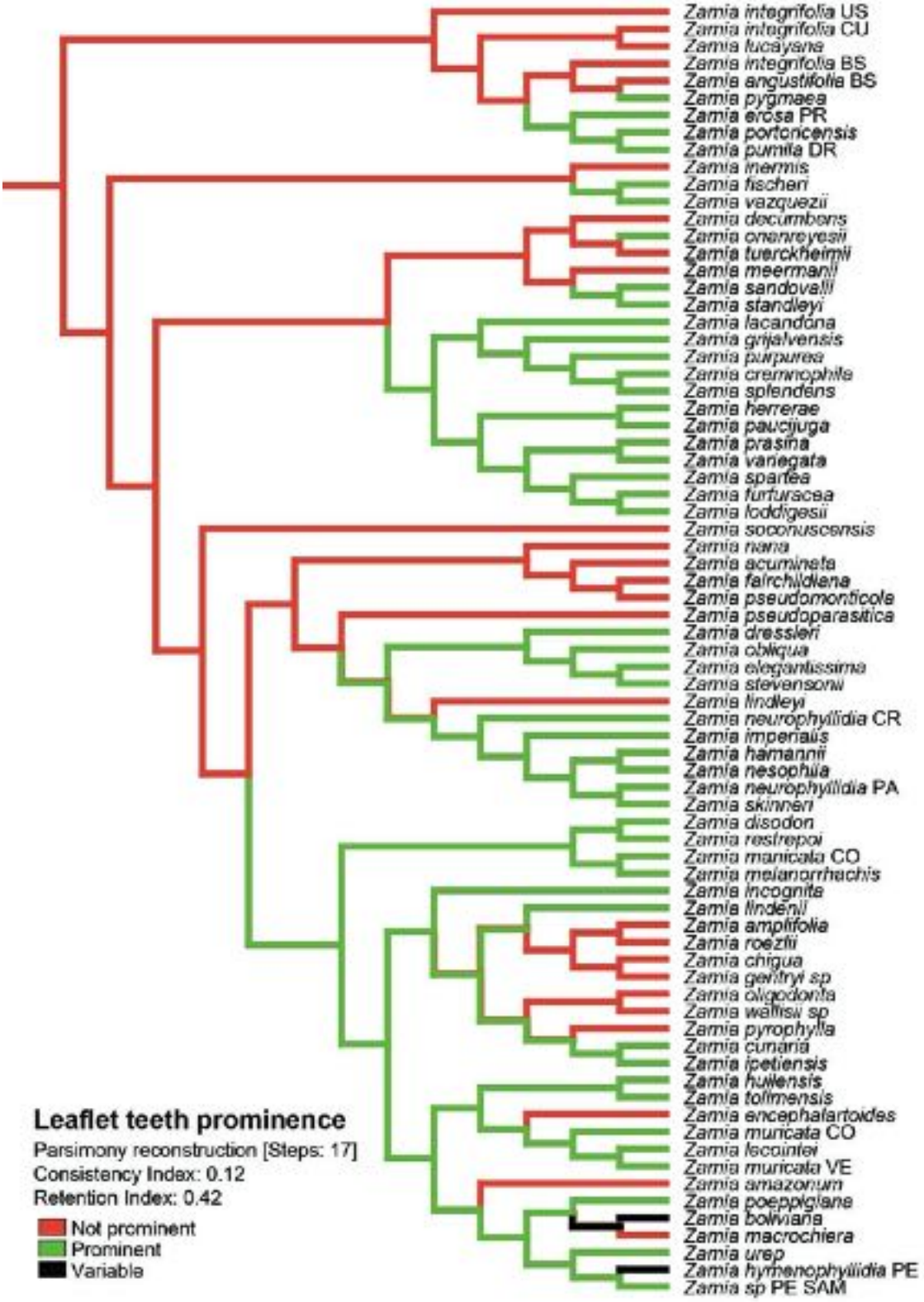
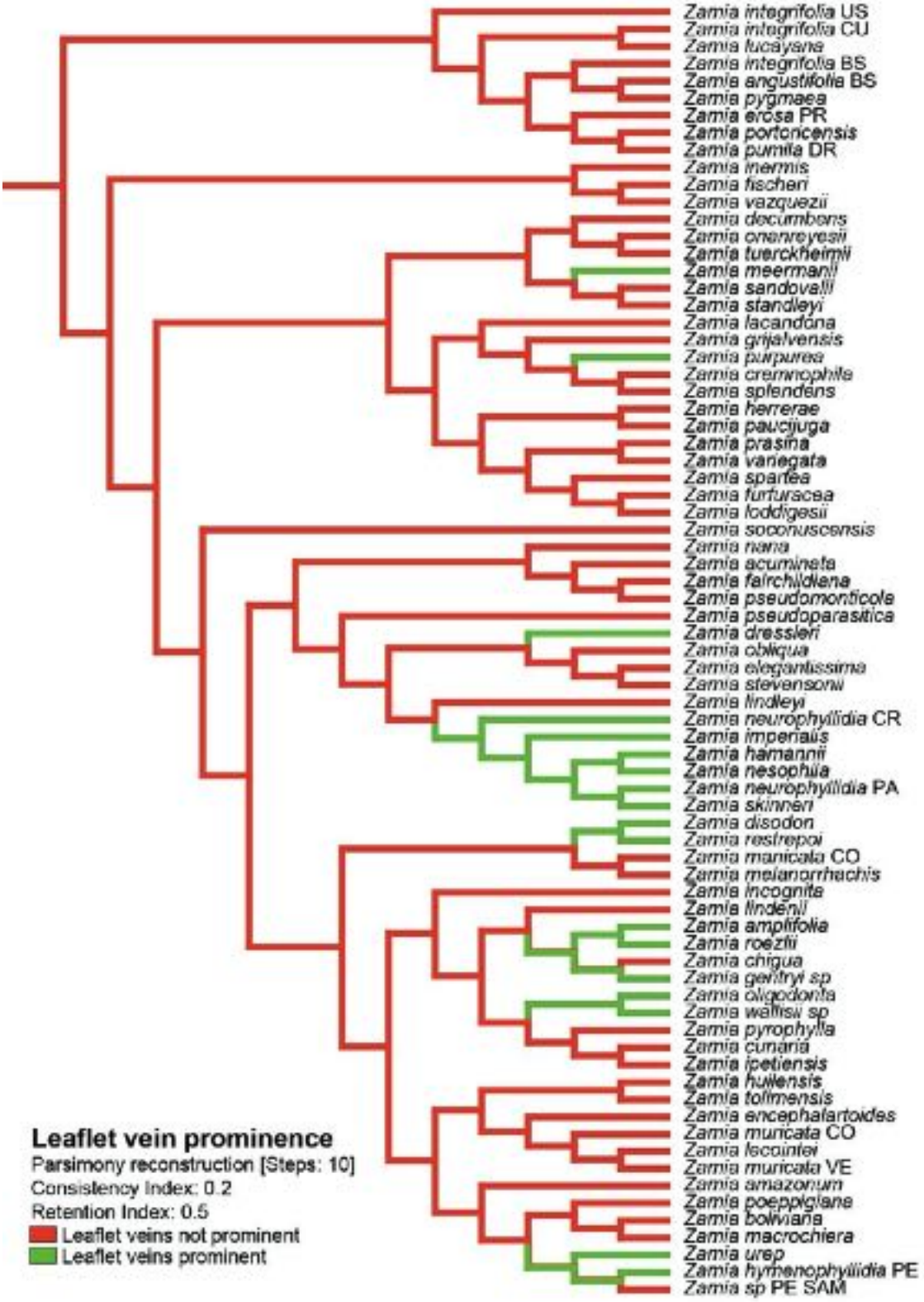
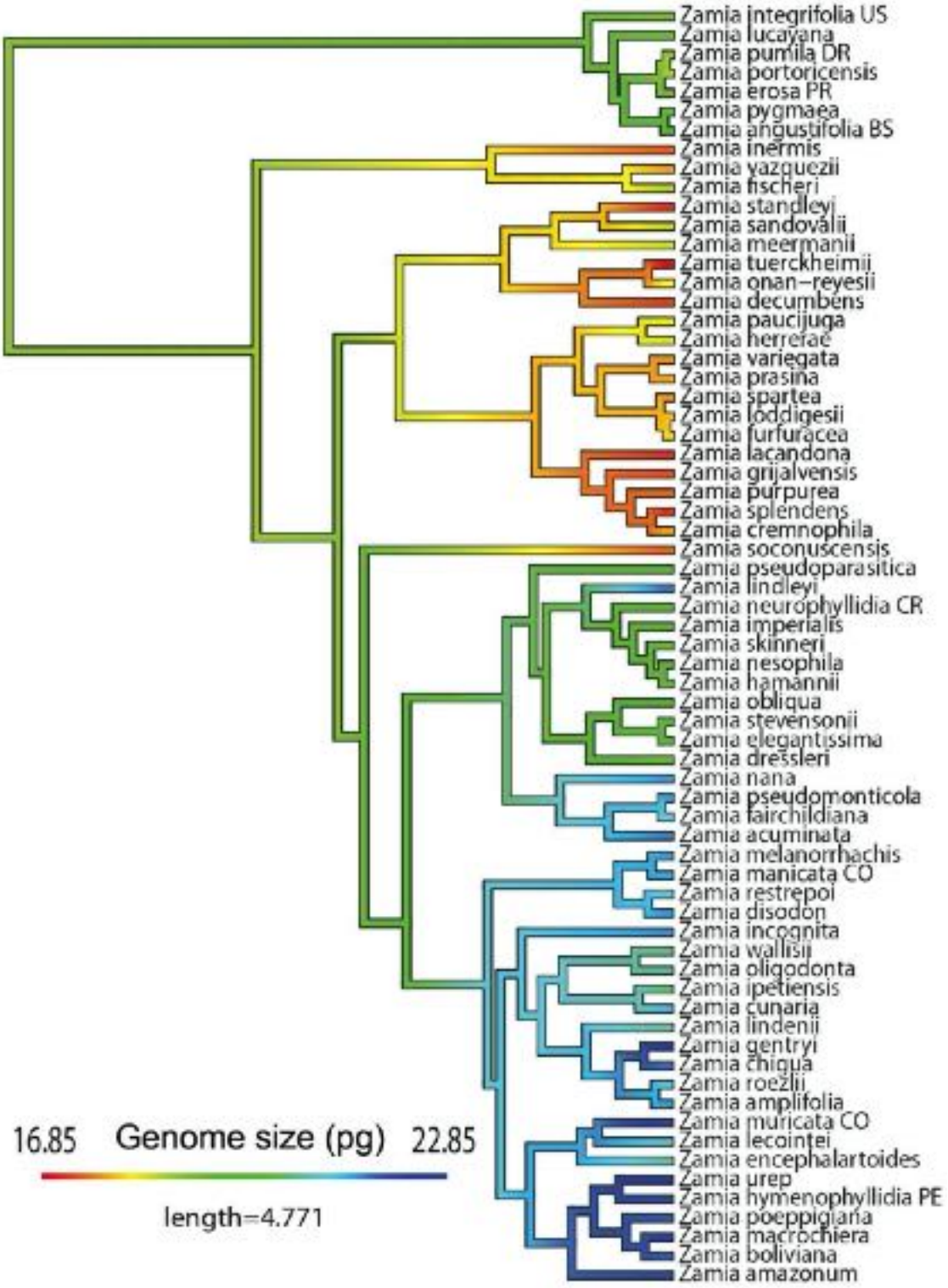


Fig. E4. Ancestral character state reconstruction for leaflet vein prominence.



APPENDIX F: GENOME SIZE IN ZAMIA AND ITS RELATIONSHIP TO PHYLOGENY

Fig. F1. Genome size in *Zamia* and its relationship to phylogeny. A strong phylogenetic signal for genome size was recovered within the crown group of *Zamia* (Pagel's lambda = 0.9721436), with both increases and decreases in genome size found throughout the genus



CHAPTER II

Population genetics of *Zamia decumbens* (Zamiaceae, Cycadales), a critically endangered cycad from the Maya Mountains of Belize

MICHAEL CALONJE BAZAR[a,b*], ALAN W. MEEROW[b,c], M. PATRICK GRIFFITH[b], JAN MEERMAN and JAVIER FRANCISCO-ORTEGA[a, d]

[a]Department of Biological Sciences, Florida International University, Miami, FL 33199, USA; [b]Montgomery Botanical Center, Coral Gables, FL 33156, USA; [c]USDA-ARS-SHRS, Miami, FL 33158, USA; [d]Fairchild Tropical Botanic Garden, Coral Gables, FL 33156, USA.

[*] For correspondence: michaelc@montgomerybotanical.org

ABSTRACT

Zamia decumbens is a critically endangered cycad species occurring in small disjunct populations on karst topography within the Maya Mountains of Belize. We examined the genetic diversity, genetic structure, and demographic history of the species using ten microsatellite loci and a variety of analytical methods. Four populations occurring in two distinct habitat types were sampled: three inside dolines (one at a cave entrance and two at the bottom of sinkholes), and one on a rocky hilltop. We found genetic variation was not structured geographically or by habitat type but rather that it reflects the demographic history and genetic connectivity of the sampled populations. Contemporary gene flow between populations is low, with a single population (Cave) facilitating most of the genetic connectivity in the region, primarily as a source of migrants to other populations. Coalescent modeling revealed that (1) the two sinkhole populations, by far the most robust known for the species, were established first, and (2) the hilltop and cave populations were more recently founded from a common ancestral population, therefore explaining their similar genetic structure despite occurring in different habitat types and at considerable geographic distance. All populations appeared generally healthy, were in Hardy-Weinberg equilibrium and were moderately heterozygous. Nevertheless, signatures for recent bottleneck events were recovered for the doline populations and a high inbreeding coefficient and average pairwise relatedness were found for the hilltop population, the latter likely the result of recent illegal harvesting activities. Ex-situ conservation efforts successfully captured a majority of the allelic diversity of the two sinkhole populations. Future efforts should focus on representing additional populations and propagating plants within existing collections in order to minimize the demand for wild collected germplasm.

Keywords: Maya civilization, ex-situ conservation, conservation genetics, gymnosperms

INTRODUCTION

The genus *Zamia* L. (Zamiaceae, Cycadales), consisting of 80 species (Calonje, Stevenson & Osborne, 2019b), is the most speciose and broadly distributed among the four neotropical cycad genera (i.e., *Ceratozamia*, *Dioon*, *Microcycas*, and *Zamia*). *Zamia decumbens* Calonje, Meerman, M.P.Griff. & Hoese (Zamiaceae, Cycadales), a species endemic to the Maya Mountains of Belize (Calonje *et al.*, 2009), belongs to a monophyletic assemblage of six species occurring in Belize, Guatemala, and northern Honduras ('Tuerckheimii clade', Calonje *et al.*, 2019a). The species occurs in small disjunct populations in lowland or submontane tropical evergreen broadleaf forest variants over hilly and steep karstic terrain (Meerman & Sabido, 2001). These karstic terrains are formed primarily over Cretaceous carbonates and to a lesser extent over Late Cretaceous to Tertiary Paleocene carbonates (Cornec, 2003, Miller, 1996). The geographic range of *Z. decumbens* is characterized by high annual rainfall ranging between 2500 to 4000 mm per year (Meerman & Sabido, 2001). However, the species occurs exclusively within two distinct habitat types that have lower water availability than the surrounding environment (Calonje *et al.*, 2009). Some populations occur on well-drained steep karstic terrains such as rocky hilltops and ridges (hereafter referred to as 'steep terrain populations'), whereas others occur inside dolines (referred to as 'doline populations'). Dolines are closed depressions occurring in a karst landscape with diameters ranging from a few meters to over 1 km, usually serving as a link between surface and underground drainage systems and typically draining into underground fissures or cave systems (Sauro, 2012). Doline populations appear generally more robust than steep terrain populations, typically hosting a higher number of plants that may attain larger dimensions and are more reproductively active. The plants are also more densely grouped in doline populations, as they are confined within the doline's perimeter and limited to the drier areas that are protected from the rain by overhanging karst ledges. The soils at the bottom of the dolines where *Z. decumbens* occurs are deep, relatively dry, and devoid of most other vegetation.

Among eight localities of *Zamia decumbens* visited during exploratory cycad fieldwork in the Maya Mountains of Belize (Calonje *et al.*, 2009), one steep terrain and three doline populations (Figs. 1, 2) were selected for population genetics studies as they were the only ones deemed large enough (i.e. $n \geq 30$) for population genetic analyses. Among the doline populations, the most robust observed were found at the

bottom of two sinkholes approximately 6.5 km distant from each other. Both sinkholes approximately 50–80 m in diameter and 30–60 m in depth and harbored nearly identical numbers of plants. One population had 178 plants (herein referred to as "Sinkhole1") and the other had 179 plants (herein referred to as "Sinkhole2"). These populations occur at the bottom of deep, steep-walled sinkholes apparently well isolated from the surrounding broad-leaved wet forest, and they contained very large plants, many of which were reproductively active during our visits. A third doline population with 52 plants (herein referred to as "Cave") was found in a very shallow depression leading to a limestone cave entrance. The Cave population, because of its location in a very shallow depression with gentle slopes, appeared much less isolated from the surrounding broad-leaved wet forest than the sinkhole populations. Finally, we found a steep terrain population of approximately 30 plants occurring on limestone ridges and hilltops south of the town of Pueblo Viejo (herein referred to as "Hilltop").

In the present study we aimed to (1) gain insight into the demographic history of *Zamia decumbens*, (2) reveal the genetic diversity, genetic structure, and gene flow patterns found among four populations of *Zamia decumbens* occurring in two very different habitat types, and (3) discuss our findings under a conservation genetics framework.

MATERIALS AND METHODS

SAMPLING, DNA EXTRACTION, GENOTYPING, AND AMPLIFICATION.

Four populations occurring along the southern terminus of the Maya Mountains in the Toledo District of Belize (Figs. 1, 2) were sampled in February 2014. Leaflets for DNA extraction were collected from all observed plants in each population to the exclusion of young seedlings carrying six or fewer leaflets per leaf. The following ten polymorphic microsatellite (SSR) loci, originally developed for Caribbean *Zamia* studies (Meerow & Nakamura, 2007, Meerow *et al.*, 2012) and previously amplified successfully from *Z. decumbens* (Griffith *et al.*, 2015, Griffith *et al.*, 2017), were selected for this study: Zam28, Zam33, Zam53, Zam59, Zam60, Zam61, Zfg23, Zfg25, Zfg32, and Zfg33. Two of these SSRs are trinucleotide repeats (Zam61 and Zfg32), and the rest are dinucleotide repeats. A total of 439 individuals were genotyped (Table 1) following previously described protocols for amplification and fragment analysis (Meerow *et al.*, 2012,

Salas-Leiva *et al.*, 2017). We used the software MICRO-CHECKER v. 2.2.3 (Van Oosterhout *et al.*, 2004) to assess stuttering patterns, large-allele dropout, and null alleles in our data. A single putative null allele was inferred at a single locus in the Sinkhole2 population; the rest of the populations showed no evidence of null alleles.

GENETIC DIVERSITY

The mean number of alleles and private alleles across all loci were calculated for every sampled population with GenAlEx v. 6.503 (Peakall & Smouse, 2012, Peakall & Smouse, 2006). Allelic richness (AR) and private allelic richness (PAR) corrected for differences in sample size (Kalinowski, 2004) were obtained with HP-Rare (Kalinowski, 2005). Expected heterozygosity (H_e) and observed heterozygosity (H_o) were calculated for all populations using Genetic Data Analysis v. 1.1. (Lewis & Zaykin, 2001). Mean pairwise relatedness (r) and the inbreeding coefficient (f) were calculated using the triadic likelihood method (Wang, 2007) as implemented in the program COANCESTRY v. 1.0.1.9 (Wang, 2011). Hardy-Weinberg equilibrium was tested for each population using global U-score tests for heterezygote excess and deficiency (Rousset & Raymond, 1995) as implemented in Genepop 4.7 (Rousset, 2008) using 10,000 Markov chain Montecarlo (MCMC) iterations. The non-random association of alleles at two loci, or linkage disequilibrium (LD) was tested for significance in each population using the likelihood ratio test (Slatkin & Excoffier, 1996) as implemented in Arlequin v. 3.5.2.2 (Excoffier & Lischer, 2010), and reported as the number of linked loci per population testing positively for LD at $P < 0.001$. Analysis of molecular variance (AMOVA) and pairwise F_{ST} values were calculated in GenAlEx utilizing 10,000 permutations. Effective population sizes (N_e) were estimated using a bias-corrected version of the linkage disequilibrium method (Waples & Do, 2008) as implemented in N_EESTIMATOR 2.0 (Do *et al.*, 2014). The results and their associated 95% confidence intervals are reported with the three default critical threshold values for the lowest allele frequency used in the analysis (0.05, 0.02, and 0.01).

ISOLATION BY DISTANCE AND BOTTLENECK ANALYSES

A Mantel test was performed in GenAlEx to determine whether the studied *Zamia decum*bens populations fit a model of isolation by distance (IBD) utilizing pairwise genetic distances (D_a; Nei et al.,

1983) between populations obtained with the program Populations v. 1.2.3.2 (Langella, 2010), and geographic distances obtained in GenAlEx using latitude and longitude coordinates in decimal degrees. Two approaches were used to assess potential bottleneck events in the studied *Zamia decumbens* populations. For one approach, a one-tailed Wilcoxon signed-rank test for heterozygosity excess under the Infinite Allele Model (Maruyama & Fuerst, 1985) was conducted as implemented in the software BOTTLENECK (Piry, Luikart & Cornuet, 1999). The program detects recent bottlenecks in a population by using the degree of heterozygosity excess across multiple loci compared to the heterozygosity expected if the population were at mutation-drift equilibrium (Cornuet & Luikart, 1996). The second approach used to explore potential bottlenecks utilized the program DIYABC v. 2.1 (Cornuet *et al.*, 2014) to calculate the ratio of the number of alleles to the range in allele size, or *M*-ratio (Garza & Williamson, 2001). The *M*-ratio is more sensitive to older, more severe bottlenecks, whereas a heterozygosity excess signal is indicative of a weaker, more recent bottleneck (Williamson-Natesan, 2005).

POPULATION GENETIC STRUCTURE

Population structure and admixture among populations was evaluated using Structure 2.3.4 (Pritchard, Stephens & Donnelly, 2000). Twenty replicates each of 1 million MCMC iterations were run (after a burn-in of 100,000) for K = 1-10 using Structure_threader (Pina-Martins *et al.*, 2017) to parallelize the runs. Using the STRUCTURE results as the input file, we utilized the web-based software StructureSelector (Li & Liu, 2018) to visualize and select the optimal number of genetic clusters using estimates of ΔK (Evanno, Regnaut & Goudet, 2005), LnP(K) (Pritchard, 2000), and the four independent estimators of Puechmaille (2016), herein referred to collectively as MMK. A membership threshold (Q-value) of 0.5 was used to assign each individual to a cluster. CLUMPAK (Kopelman *et al.*, 2015) as implemented in StructureSelector (Li & Liu, 2018) was used to summarize the independent runs from each K and to visualize the optimal alignment of inferred clusters in a histogram labeled by population.

PHYLOGENETIC PLACEMENT AND DEMOGRAPHIC HISTORY

We examined the phylogenetic positioning and historical demography of *Zamia decumbens* utilizing the time-calibrated phylogeny of Calonje *et al.* (2019a). We also used approximate Bayesian computation

as implemented in the program DIYABC v. 2.1 (Cornuet *et al.*, 2014) to compare competing evolutionary scenarios and estimate their relative likelihood. A total of five invasion and five vicariance scenarios were simulated (Figs. A1 & A2) as described in Meerow *et al.* (2012), Salas-Leiva *et al.* (2017), and Meerow *et al.* (2018). Statistics used in the simulation scenarios were mean genic diversity and the *M* coefficient (Garza & Williamson, 2001) for single samples, and shared allele distance, F_{ST}, and index of classification (Rannala & Mountain, 1997, Pascual *et al.*, 2007) for pairwise comparisons. Four million simulations were run using default priors except for maximum mutation rate which was set at 5.04×10^{-4} following Meerow *et al.* (2018). The confidence in scenario choice was evaluated using the direct approach on the 500 simulated datasets closest to the observed dataset and the logistic approach on the closest 50,000 datasets. The prior and posterior predictive error results were summarized with standard statistics.

NETWORK AND CONTEMPORARY GENE FLOW ANALYSES

A network analysis was conducted in EDENetworks (Kivela, Arnaud-Haond & Saramaki, 2015) using F_{ST} values as described by (Salas-Leiva *et al.*, 2017) and allowing the program to automatically determine the distance (i.e., threshold) beyond which links in the network are removed. The robustness of the resolved network was evaluated by bootstrapping.

We assessed contemporary gene flow between populations of *Zamia decumbens* utilizing the program Bayesass 3.0.4 (Wilson & Rannala, 2003), which uses a coalescent approach to estimate pairwise recent immigration rates between populations. Samples were run for 3×10^7 MCMC iterations with a burn-in of 1×10^7 and sampling every 2000 generations. Convergence of the MCMC analysis was evaluated using Tracer v1.6 (Rambaut *et al.*, 2014) and deemed acceptable based on examination of the trace file and the high ESS scores for all parameters. Four additional independent runs with different seeds were performed, and the posterior mean parameters were found to be concordant with the results of the first run. We used the software GeneClass2 (Piry *et al.*, 2004) to identify putative first-generation migrants following the Bayesian criterion of Paetkau *et al.* (1995) and the simulation algorithm of Paetkau *et al.* (2004) with 1000 simulated individuals, an alpha level of 0.01, and a likelihood ratio L_{home}/L_{max}.

RESULTS

GENETIC DIVERSITY

The 10 SSR loci were polymorphic across all populations and uniquely genotyped 426 out of 439 individuals (97.0%); for Hilltop, 30 out of 30 (100%); for Cave, 52 out of 52 (100%); for Sinkhole1, 176 out of 178 (98.9%); and for Sinkhole2, 168 out of 179 (93.9%). The average number of alleles per locus ranged from 2.75 (Zfg32) to 15.5 (Zam53) (Table 2). The maximum mean number of alleles per site was found in Cave (Na = 7.4) and the minimum in Sinkhole1 (Na = 4.0). When corrected for sample size, allelic richness was similarly highest in Cave (4.09) and lowest in Sinkhole1 (2.89). The highest H_o and H_e (Table 3) were found in Cave (H_o=0.72, H_e=0.70), and the lowest were found in Hilltop (H_o=0.53, H_e=0.54). None of the populations had significant heterozygote deficiency or excess in the (U) tests, indicating they are likely in Hardy-Weinberg equilibrium. The highest number of private alleles was found in Sinkhole2 (14), followed by Cave (11), Hilltop (5), and Sinkhole1 (2). The Cave population had the highest PAR value (0.95), followed by Sinkhole 2 (0.69), Hilltop (0.62), and Sinkhole1 (0.55). The genetic variation across the 10 SSRs (Table 4), determined by permuted AMOVA, was 21% among populations, 79% within populations. Permuted mean F_{ST} was 0.207 ($P < 0.0001$). Pairwise permuted F_{ST} values (Table 5) were moderate to high, ranging from 0.103 (Hilltop-Cave) to 0.268 (Hilltop-Sinkhole1). The greatest differentiation was observed between Hilltop and both Sinkhole1 (0.268) and Sinkhole2 (0.241) populations, the lowest between the Cave population and both Hilltop (0.103) and Sinkhole2 (0.116) populations. Mean pairwise relatedness (Table 3; Wang, 2007) was highest in Hilltop (0.416) and Sinkhole1 (0.355) and lowest in Sinkhole2 (0.25) and Cave (0.195). The inbreeding coefficient followed a similar pattern, being highest in Hilltop (0.31), followed by Sinkhole1 (0.20), Sinkhole2 (0.18) and Cave (0.11). Significant linkage disequilibrium between loci was detected in all populations, affecting between 11.1% and 62.2% of the pairs of loci in each population (Table 2).

ISOLATION BY DISTANCE AND BOTTLENECK ANALYSES

The results of the Mantel test showed no correlation between genetic divergence and geographic distances, indicating no IBD (R_2 = 0.0023, R_{xy} = 0.048, P= 0.310). The one-tailed Wilcoxon tests identified

the Cave ($P < 0.001$), Sinkhole1 ($P < 0.001$), and Sinkhole2 ($P < 0.05$) populations as having a significant heterozygosity excess as a consequence of a recent genetic bottleneck, while the *M* test did not detect a significant reduction ($M < 0.7$) in population size in any of the *Zamia decumbens* populations (Table 3).

POPULATION GENETIC STRUCTURE

The admixture model (with allele frequencies correlated) was the best fit for the SSR data. The *K* estimators resolved different optimal numbers of distinct genetic populations: ΔK found K=2 as optimal, LnP(K) found K=10, and the MMK statistics were split between K=3 and K=4. If K=2 is visualized, Sinkhole1 is clearly distinct from the rest of the populations, whereas in K=3, Hilltop and Cave are assigned to the same cluster, and each sinkhole population belongs to its own distinct cluster with Sinkhole2 being slightly more admixed than Sinkhole1 (Fig. 3).

DEMOGRAPHIC HISTORY

Model checking analyses in DIYABC indicated a good fit between the scenarios and the posterior distributions of the analysis parameters for the SSR data. Of the ten scenarios analyzed (Figs. A1 & A2), scenario 2 of the invasion model (Fig. 4) had the highest posterior probability (PP). In scenario 2, the two sinkhole populations were the first to diverge, followed by the Cave and Hilltop populations. The confidence in the scenario choice is high with posterior predictive errors (i.e., proportion of wrongly identified scenarios over 1000 test data sets) being 0.093 and 0.067, and the prior predictive errors being 0.182 and 0.165 using the direct and logistic approaches, respectively. The type I error, or the proportion of times in 1000 test datasets that scenario 2 did not have the highest posterior probability among the two highest scoring scenarios (scenarios 2 and 4) was 0.067 in the logistic regression and 0.069 using the direct approach. The type II error, or the proportion of times in 1000 test datasets that scenario 2 has the highest posterior probability among the 5 invasion scenarios was 0.072 in the logistic regression and 0.095 using the direct approach. Following the scenario in Fig. 4, DIYABC estimated a mean mutation rate (μmic) for *Zamia decumbens* of 3.01×10^{-4} per generation (Table 6). The Hilltop population had the highest N_e (259.2), followed by Sinkhole2 (47.5), Sinkhole1 (43.5), and Cave (34.5) (Table 7). The NeEstimator results, calculated using the LD method, were comparable albeit with Cave having slightly higher N_e than

Sinkhole1 (Table 8). The coalescence time between Hilltop and Cave (t1 in Fig. 4) was estimated by DIYABC at between 77.8 (30-year generation time) and 155.6 (60-year generation time) Kyr ago (KYA). For Sinkhole1 and the lineage leading to Cave and Hilltop (t2 in Fig. 4), coalescence time was estimated at between 185.5 (30-year generation time) and 371.0 KYA (60-year generation time). Lastly, the coalescence time for Sinkhole2 and the lineage including the other three populations (t3 in Fig. 5) was estimated at between 191.1 (30-year generation time) and 382 KYA (60-year generation time).

NETWORK AND CONTEMPORARY GENE FLOW ANALYSES

The automatically thresholded (0.23) F_{ST} network by population of *Zamia decumbens* SSR genotypes (Fig. 4) appeared robust by the bootstrap tests (not shown). The Cave population had the largest number of connecting edges and the greatest betweenness centrality, as the three other populations had edges connecting exclusively to it. The Bayesass analysis results show that all populations were generally self-seeding, with 96 to 99 percent of individuals originating from within the same site (Table 9). Very low levels of contemporary gene flow (i.e., proportion of individuals within a population that are migrants) between populations were found, with mean estimates ranging from 0.002 to 0.016. The differences between these estimates were not significant, as the 95% HPD intervals all overlapped. The GeneClass2 analysis identified a total of 10 putative first-generation migrants ($P < 0.01$), with most migrants originating from the Cave population. Seven individuals from Sinkhole2 and two individuals from Hilltop were identified as migrants from Cave, and one individual from Cave was identified as a migrant from Hilltop (Table 10).

DISCUSSION

PHYLOGENETIC PLACEMENT AND DEMOGRAPHIC HISTORY OF ZAMIA DECUMBENS

According to a time-calibrated species tree phylogeny of the genus *Zamia* that used a multilocus sequence dataset of 10 independent loci (Calonje *et al.*, 2019a), *Z. decumbens* belongs to a well-supported clade (Tuerckheimii clade; posterior probability [PP] = 0.96) of six rainforest-dwelling species occurring in adjacent areas of Guatemala, Belize, and Honduras (Fig. 6, Calonje *et al.*, 2019a). The mean crown age

estimated for the Tuerckheimii clade was 2.39 Ma, with most of the cladogenic events occurring during the Pleistocene (Calonje et al., 2019a). The 95% highest posterior density (HPD) for the estimate is broad (0.15–3.07 Ma) and includes part of the Pliocene. However, a Pleistocene origin for the clade is more likely, as palynological evidence from regional fossil floras indicates that wet rainforest environments such as those currently inhabited by all members of the Tuerckheimii clade were either not present or were poorly represented in the region during the Pliocene (Graham, 1998). Within the Tuerckheimii clade, *Z. decumbens* is the earliest branching species of a subclade of three large arborescent species (Fig. 6, Calonje *et al.*, 2019a) that originated in the Pleistocene approximately 1.25 Ma (95% HPD 0–1.94).

As *Zamia decumbens* occurs exclusively on karst landscapes in the Maya Mountains, its origin is likely associated with the karstification of the ancient (Cretaceous to Paleocene) carbonate surfaces on which it occurs. The species occurs in four separate karst regions in Belize (*sensu* Miller, 1996): K-T Fault Ridges, Little Quartz Ridge, Vaca Plateau, and Boundary Fault. The Cave, Sinkhole1 and Sinkhole2 populations occur in the Little Quartz Ridge area, presumably formed by Late Cretaceous (100.5–66 Ma) limestones of the Campur formation (Miller, 1996), whereas the Hilltop population occurs in the K-T Fault Ridges, assigned by Cornec (2003) to the La Cumbre Formation of Late Cretaceous (100.5–66 Ma) to Paleocene age (66–56 Ma). The K-T Fault Ridges where the Hilltop population occurs consists of a series of isolated block-fault ridges that trend southwest to northeast and are separated from the Little Quartz Ridge karst region by plains of the Toledo Series (Miller, 1996), a mainly Paleogene suite of conglomerates, mudstones, siltstones, and sandstones (Fisher & King Jr, 2015). While the carbonates underlying *Z. decumbens* are quite old, the karst landscapes in which the species occurs are of more recent origin.

Karstification in the southern Vaca Plateau, one of the karst regions in which *Zamia decumbens* occurs, began ca. 700 ka (Miller, 1990). Closed depressions such as those hosting the largest and most robust populations of *Z. decumbens* are common in Belize karsts, and they are likely collapse features sited over or near cave networks (Miller, 1996). Consequently, the establishment of these populations is likely linked closely to the formation of caves in the region. The development of caves in the karsts of the southern Vaca Plateau began as early as 176 ka and was well advanced by 102 ka (Miller, 1996). These

dates are congruent with the findings of the DIYABC analysis, which estimated a coalescence date for the populations studied of between 191.1 to 95.5 ka, using generation time estimates of 30 and 15 years, respectively (Table 7).

Various lines of evidence detailed above, including palynological, geological, and molecular evidence suggest a Pleistocene origin for *Zamia decumbens*. In fact, the Pleistocene appears to be the most important period in the evolution of extant Neotropical cycad species as a whole, with independent phylogenetic studies showing the majority of cladogenic events occurred during this time for *Ceratozamia* (Medina-Villarreal, González-Astorga & Espinosa de los Monteros, 2019), *Dioon* (Dorsey *et al.*, 2018), and *Zamia* (Calonje *et al.*, 2019a). The Pleistocene, lasting from about 2.58 Ma to 12.0 ka (Walker *et al.*, 2018) was a geological epoch characterized by intermittent cycles of warming and cooling, with temperature and humidity generally increasing during glacial minima, and decreasing during glacial maxima. During glacial maxima, temperatures were about 6° C cooler than today, and tropical rainforests contracted, while arid savannas and dry forest ecosystems expanded, whereas the opposite occurred during glacial minima (Bridgewater, 2012). Indeed, palynological and other fossil evidence from lake sediments suggests that some areas in Belize currently hosting tropical rainforests supported dry and open vegetation during glacial maxima (Bridgewater, 2012, Larmon *et al.*, 2019). The southernmost Maya Mountains region, hosting the largest and most robust populations of *Zamia decumbens* known, was considered to be a primary Pleistocene refuge by Toledo (1982) on the basis of climatic and topographic data, as well as patterns of species endemism, richness, and geographic distribution. If the southern Maya Mountains were indeed a Pleistocene refuge, it may have allowed moist forests to persist during arid climate phases, ensuring the survival of *Zamia decumbens* despite the significant climate fluctuations of the Pleistocene.

THE MAYA CIVILIZATION: HISTORIC DEFORESTATION IN THE REGION AND RITUAL USE OF CAVES AND SINKHOLES

Several Mayan archaeological sites exist within the distribution range of *Zamia decumbens*, the most important of these being the city of Caracol, which is located on the Vaca Plateau at the western edge of the Maya Mountains. It was one of the most important Mayan political centers during the Classic Period,

having an area of continuous settlement extending over 200 sq. km (Chase *et al.*, 2011) and hosting at least 100,000 people (Chase & Chase, 1994) at its height between AD 600 and 700 (Chase & Chase, 2017). The population of Caracol at its peak was higher than present-day Belize City, currently the largest city in Belize. A vegetation reconstruction using $\delta^{13}C$ values of fulvic acids extracted from cave sediments at Reflection Cave on the Vaca Plateau (Polk, van Beynen & Reeder, 2007) indicated substantial environmental changes such as widespread agriculture-driven deforestation and cyclical droughts occurred during the Late Holocene from ca. 2600 cal yr BP to ca. 1500 cal yr BP. Although most known populations of *Z. decumbens* occur within protected areas in areas that today are relatively well-forested, the evidence of widespread deforestation in the region associated with Caracol suggests that these forests may have experienced severe disturbances a few centuries ago. Even so, it is unclear whether potential deforestation would have affected the steep hilltops and dolines where *Z. decumbens* occurs, as they are seemingly unsuitable for agriculture. It is well documented that cycads in Mexico and Northern Central America (MNCA) have traditionally been used for millennia as a source of food starch obtained from seeds or stems (Bonta *et al.*, 2019). Most typically in MCNA, the much larger seeds of *Dioon* and *Ceratozamia* were used for food, and to a lesser extent the stems of some *Zamia* species were also used. It is unknown whether the Maya peoples utilized *Z. decumbens* as a food source, but they most certainly were aware of its existence, as the Mayans commonly used cave entrances and sinkholes for ritual purposes. A study of cave entrances and sinkholes associated with the Mayan city of Caracol (Ishihara-Brito, Awe & Chase, 2011) found that almost all accessible cave entrances had evidence of ritual use, including speleothem breakage, architectural modifications, petroglyphs, and artifacts. During the course of our fieldwork, we observed the remains of a charcoal kiln as well as pottery shards at the bottom of the sinkhole hosting the Sinkhole1 population. These shards may be of Late Classic types (A.D. 600-900) as were the majority of those observed in caves entrances associated with Caracol (Ishihara-Brito *et al.*, 2011). Although it is unclear whether any of the individual plants observed during our fieldwork would have been alive before the Maya peoples abandoned settlements in the region over 500 years ago, the estimated coalescence dates of the studied populations (Table 7) suggests that these populations were all founded prior to the establishment of the Maya civilization. Additional research such as the analysis of starch grains associated with grinding

implements or palynological studies of cave and sinkhole sediments may bring further insight into whether the Maya civilization may have had an effect on doline populations either by direct harvesting of plants or by habitat modification for ritual purposes.

POPULATION GENETIC STRUCTURE, DIFFERENTIATION, AND GENE FLOW BETWEEN POPULATIONS

The three doline populations (Sinkhole1, Sinkhole2, and Cave) are closest to each other geographically, occupy similar habitats in the same karst region, and are separated from the Hilltop population by the non-carbonate plains of the Toledo Series (Fig. 1). Notwithstanding the habitat similarities and geographical proximity of the doline populations, these did not form a distinct genetic group separate from the distant Hilltop population as expected. Instead, Cave and Hilltop were the two least differentiated populations from each other (Fig. 3, Table 5) despite occurring in different habitat types and separated by a considerable 15 km distance. The geographic distance between the studied populations does not appear to be a major driver in shaping their genetic divergence, as no pattern of isolation by distance was detected by the Mantel test. As contemporary gene flow between all sampled populations is very low (Table 9), the similarities in genetic structure between the Cave and Hilltop populations may best be explained by their establishment from a common ancestral population via dispersal, as suggested in the optimal scenario selected in the DIYABC analysis (Fig. 4).

The Bayesass analysis results show that although contemporary migration between populations is rare, it does appear to occur, as none of the 95% confidence intervals for migration between populations overlapped zero. Moreover, a limited amount of migration between some populations appears likely as evidenced by the detection of ten potential first-generation migrants in the Geneclass2 analysis (Table 10). The relative contribution of each population to gene flow in the region varies widely, with the Cave population identified as the most important in maintaining genetic connectivity in two separate analyses. The EDENetworks network analysis results identified the Cave population as the single population exclusively connecting all populations together, and the Geneclass2 results identified the Cave population as either a source or destination of all putative first-generation migrants (Table 10). However, it is

important to note that the gene flow involving the Cave population appears highly directional, with all but one of the potential migrants originating in the Cave population and only a single migrant received by it from Hilltop. The fact that no potential migrants originated in either sinkhole population yet seven were received by Sinkhole1 suggests that the sinkhole topography, characterized by considerable depth and steep vertical walls most readily allows migration into the sinkholes but appear to present a formidable barrier to outward migration.

Although a limited amount of contemporary gene flow between populations does appear to exist, little is known about the biological mechanisms underlying gene dispersal between populations. Cycads are exclusively dioecious, reproducing by means of pollen and seed producing strobili borne on separate plants. Most species are believed to be insect-pollinated, and in most species the pollination agents are beetles (Terry *et al.*, 2012). Two major lineages of beetles inhabit New World cycad cones: beetles in the family Erotylidae are found in all New World genera, whereas weevils in the subtribe Allocorynina are restricted to *Dioon* and *Zamia* (Tang *et al.*, 2018, O'Brien & Tang, 2015). Two putative pollinators have been observed feeding and reproducing in dehiscent pollen cones of *Z. decumbens*: an unidentified erotylid beetle in the genus *Pharaxonotha* Reit., and the weevil *Rhopalotria calonjei* Tang & O'Brien (O'Brien & Tang, 2015). Pollination of cycads appears to occur within a very restricted geographic range (Yang & Meerow, 1996), and seed dispersal capabilities are similarly limited (Hall & Walter, 2013, Tang, 1989, Tang, 1993). Populations of the cycad *Ceratozamia* aff. *latifolia* in San Luis Potosí, Mexico have a remarkably similar distribution pattern to *Z. decumbens*, with individuals occurring at greatest concentrations near cave entrances and sinkholes (Hoese, 2012). Hoese (2012) suggests fruit bats as a putative long-distance dispersal agent for seeds of *C.* aff. *latifolia* on the basis of several observations including the location of the cycad populations near probable bat habitats, and the ability of the bats to carry fruits for long distances between roosts. Bat-mediated dispersal is also a possibility for *Z. decumbens*, as it inhabits similar localities near potential bat habitats. However, although in both sinkhole habitats we observed scratch marks on seed sclerotestas suggesting animals had fed on the sarcotestas, we were not able to ascertain the identity of the potential dispersal agent(s). The identification of the seed dispersal

agents of *Z. decumbens*, along with an understanding of their travel patterns would help explain the interesting distribution patterns observed in the species and should be a focus of future studies.

THE REMARKABLE ROBUSTNESS OF SINKHOLE POPULATIONS

The two sinkhole populations are by far the most robust known populations of *Zamia decumbens.* Each sinkhole hosts approximately 180 plants each, making them the largest known populations of the species. The plants in the sinkhole populations attain larger dimensions and are more often reproductively active than those observed in other habitat types. Moreover, the plants are heavily concentrated within a small area, as they are surrounded by the steep sinkhole walls that limit their outward dispersal, and within the sinkhole itself they are limited to the drier areas below overhangs, where they are protected from direct rain.

The uniqueness of the sinkhole habitat and the robustness of the *Zamia decumbens* populations found within has earned the species the common name the 'sinkhole cycad'. However, this habitat preference appears to be shared by at least one other cycad species. Hoese (2012) reported a similar distribution pattern for *Ceratozamia* aff. *latifolia* in San Luis Potosí, Mexico to what we observed in *Z. decumbens*: a greater concentration of plants occurring near cave entrances and within sinkholes, and smaller local populations occurring elsewhere in the karst landscape. Clearly, several characteristics associated with sinkhole habitats in a karst environment appear conducive to supporting large, healthy cycad populations.

The drier soils in the rainfall-protected areas of the sinkholes appear to provide a competitive advantage to cycads such as *Zamia decumbens* which are more drought-tolerant than most rainforest-adapted vegetation and are therefore able to thrive here with little competition. In fact, in both sinkholes, *Z. decumbens* was the dominant species in these dry zones, with very few other plant species present. In addition to being relatively dry, the soils associated with doline populations are generally much deeper than those found in the steep terrain habitats, and likely more fertile overall judging by the large guano deposits we observed on the sinkhole floors.

The small size of the sinkholes combined with their significant depth which appears to impede outward migration of seeds has resulted in an extremely high population density at both sinkhole

populations. A high concentration of plants likely makes pollination more efficient in these habitats than in habitats where plants are more scattered from each other such as in the Hilltop population.

Sinkholes and other topographical depressions provide vegetation with a stable environment in which fluctuations in humidity and temperature are minimized, and where they are protected from many hazards including desiccating wind, excessive solar radiation, and fires (Hardin, 1954, Hoese, 2012).

Hoese (2012) noted the larger size of *Ceratozamia* aff. *latifolia* populations occurring within sinkholes than in the surrounding karst landscape and attributed the size difference to the age of establishment of the populations. According to Hoese, the sinkholes serve as refuges in times of high environmental stress, allowing these populations to survive when surface populations may not be able to, and allowing them to serve as a source of germplasm to establish new colonies. The DIYABC analysis results (Table 7) are consistent with Hoese's (2012) antiquity hypothesis, as it showed that the two sinkhole populations were established much earlier than both the Cave and Hilltop populations. The sinkhole populations were by far the most reproductively active populations we encountered, and thus it would seem that they could serve as germplasm sources to found new populations during times of environmental stress, as suggested by Hoese (2012). However, the extremely low outgoing migration rates reported for the sinkhole populations in the Bayesass analysis (Table 9), and the lack of putative first-generation migrants deriving from either sinkhole population detected in the Geneclass2 analysis (Table 10) indicate that germplasm dispersal outside of the deep sinkholes is likely very rare. However, populations occurring in shallower dolines, such as the Cave population, may still be buffered enough against environmental changes enabling them to survive in the long term, and may be a more likely source of germplasm to found new populations, as suggested by its high genetic connectivity (Fig. 5) and the nine putative first-generation migrants that originated there according to the Geneclass2 analysis (Table 10).

CONSERVATION GENETICS

Zamia decumbens is considered a Critically Endangered species because of the fragmented nature of its distribution, the small number of known populations, the small combined area of occupancy for the species, and observed evidence of past and present commercial exploitation of plants and seeds (Calonje *et*

al., 2009). The species is currently included in the IUCN Redlist, albeit under the misapplied name *Z. prasina* (Stevenson, 2010), a name which correctly refers to a different species with a widespread distribution in drier areas of Mexico, Guatemala and Belize (Calonje & Meerman, 2009). Despite its Critically Endangered status, the individual populations visited occurred in protected areas (Cave and SH2 in Columbia Forest Reserve) or little disturbed habitats (SH1 and Hilltop) and appeared generally healthy, with ample evidence of reproduction and recruitment. All populations were in Hardy-Weinberg equilibrium, retained moderate levels of heterozygosity (Table 3), and were moderately to highly differentiated from each other (Table 5).

The three doline populations were the only populations to exhibit a signature of recent bottlenecks with the Wilcoxon tests combined with high levels of linkage disequilibrium (LD, Table 3), suggesting the LD may have been caused by the recent bottlenecks experienced by these populations. The increase in LD can occur as a result of the loss of some haplotypes after a bottleneck and can be exacerbated by the increasing effects of genetic drift after a period of small population size (Slatkin, 2008). Levels of inbreeding and pairwise relatedness were high in most populations (Table 3), and highest in the Hilltop population. The Hilltop population hosted smaller plants and appeared less reproductively active than the three doline populations. According to our local field guide, the Hilltop populations had been a source of commercial extraction of plants and seeds in the recent past, with the largest, most reproductively active plants being preferentially harvested. The reported prior removal of large adult plants from this population may have resulted in a higher mean inbreeding coefficient and average pairwise relatedness compared to other populations, as our sampling may have been biased toward younger plants of similar age that are likely to include a higher proportion of siblings than an undisturbed population.

The strong to moderate genetic differentiation between populations and the small migration rate indicates most populations are genetically isolated, and this may be one of the causes for the high levels of inbreeding, mean pairwise relatedness and linkage disequilibrium prevalent in these populations. Even so, the populations have been able to retain high levels of genetic diversity (heterozygosity and allele richness) despite the low levels of contemporary gene flow between them. Perhaps the long generation times and presence of multiple overlapping generations prevents a detectable genetic signal (i.e. loss of

heterozygosity) from arising for many years. Additionally, the dioecious reproduction of cycads may help preserve heterozygosity as it prevents self-fertilization.

The small and fragmented nature of populations of *Zamia decumbens*, coupled with their genetic isolation means they are highly susceptible to the effects of genetic drift and inbreeding, or to potential catastrophic events. The doline populations, although possibly more protected from environmental changes affecting the surrounding forests than steep terrain populations, are especially vulnerable to catastrophic events as a consequence of the small geographic area they cover and their isolation from neighboring populations. Conversely, populations in steep terrain environments such as ridgetops may be more vulnerable to environmental changes but are typically not as geographically isolated as doline populations and so may be more resilient to catastrophes affecting small geographic areas, such as landslides or localized fires.

As part of our fieldwork with *Zamia decumbens*, we also collected seeds for cultivation in ex-situ collections from the two sinkhole populations which were the only ones with mature seeds at the time. A previous study found that a collection of three seed cones from Sinkhole1 was able to capture 70% of the allelic diversity of the populations, and a collection of four seed cones from Sinkhole2 captured 73.6% of the diversity (Griffith *et al.*, 2015). These two populations are now reasonably well represented in ex-situ conservation, but neither the Cave nor the Hilltop populations are represented. Future germplasm collection efforts should focus on previously unsampled populations in order to better represent and preserve the genetic diversity of this remarkable species; and living plant collections at botanical gardens should be actively propagated to make this species more available in horticulture and thus minimize horticultural demand for wild-collected plants. Similar efforts in botanical garden propagation of the Cuban cycad *Microcycas calocoma* have successfully reduced demand for the species as exemplified by the reduction in the price of seeds over time (Kay, 2011).

CONCLUSIONS

Zamia decumbens belongs to a clade of rainforest dwelling species that appears to have primarily diversified during the Pleistocene. Its demographic history appears closely linked to the karstification of the limestone bedrock on which it occurs as well as with the formation of caves necessary to form the collapse features (i.e., sinkholes) in which the largest, most reproductively active known populations occur. The two sinkhole populations studied here are the largest known population of this species, and our results suggest that they were the first to be established. We found that the genetic variation of the studied populations was structured neither geographically nor by habitat type, but rather seemed to reflect the demographic history of the populations and their genetic connectivity. The latter appears to be largely dependent on the Cave population, which serves as a hub genetically linking all other populations, serving primarily as a source of first-generation migrants to other populations. Future research should focus on identifying the seed dispersal agents for the seeds of this species in hopes of gaining a better understanding of the reasons behind the unusual genetic variation patterns found in this species.

Despite occurring in small isolated populations within the Maya Mountains and considered Critically Endangered, most known populations of *Z. decumbens* occur in protected areas, and the populations surveyed appear reasonably healthy. All populations were in Hardy-Weinberg equilibrium, moderately heterozygous, and did not exhibit decreased heterozygosity. Signatures for recent bottleneck events, however, were recovered for the doline populations. Anecdotal evidence of commercial extraction of mature plants from the Hilltop population appears to be corroborated by the high average pairwise relatedness and inbreeding values recovered for this population in the COANCESTRY analysis. Ex-situ conservation efforts have successfully captured a majority of the allelic diversity of the two sinkhole populations. Future efforts should focus on representing other populations in ex-situ collections and in the propagation of plants in existing collections to make seeds available to the horticultural industry and therefore minimize the demand for wild-collected germplasm. The establishment of local inter-situ sustainable management nurseries for propagating cycads should also be explored, as this model has been previously implemented in Mexico with some success in protecting cycad habitats, discouraging illegal collecting, and benefiting local villages (Vovides, Pérez-Farrera & Iglesias, 2010).

ACKNOWLEDGMENTS

We are grateful to the Belize Forest Department, Belize Defense Force, Belize Botanic Gardens, Green Hills Botanical Collections, Valentino Tzub, Geoffrey Hoese, Chad Husby, Freddy Tut, and the Ya'axché Conservation Trust for help with permitting, logistics, and fieldwork. We would also like to thank Kyoko Nakamura, Dayana Salas-Leiva, and Vanessa Sanchez for training and assistance with microsatellite lab work.

LIST OF REFERENCES

Bonta, M, Pulido-Silva, MT, Diego-Vargas, T, Vite-Reyes, A, Vovides, AP, Cibrián-Jaramillo, A. 2019. Ethnobotany of Mexican and northern Central American cycads (Zamiaceae). *Journal of Ethnobiology and Ethnomedicine* **15:** 4.

Bridgewater, S 2012. *A Natural History of Belize : Inside the Maya Forest*: University of Texas Press.

Calonje, M, Meerman, J. 2009. What is *Zamia prasina* (Zamiaceae: Cycadales)? *Journal of the Botanical Research Institute of Texas* **3:** 43-49.

Calonje, M, Meerman, J, Griffith, MP, Hoese, G. 2009. A new species of *Zamia* (Zamiaceae) from the Maya Mountains of Belize. *Journal of the Botanical Research Institute of Texas* **3:** 31-41.

Calonje, M, Meerow, AW, Griffith, MP, Salas-Leiva, D, Vovides, AP, Coiro, M, Francisco-Ortega, J. 2019a. A Time-Calibrated Species Tree Phylogeny of the New World Cycad Genus *Zamia* L. (Zamiaceae, Cycadales). *International Journal of Plant Sciences*.

Calonje, M, Stevenson, DW, Osborne, R. 2019b. The World List of Cycads, online edition (http://cycadlist.org) Accessed (Accessed 21 Sep 2019).

Chase, AF, Chase, DZ, Weishampel, JF, Drake, JB, Shrestha, RL, Slatton, KC, Awe, JJ, Carter, WE. 2011. Airborne LiDAR, archaeology, and the ancient Maya landscape at Caracol, Belize. *Journal of Archaeological Science* **38:** 387-398.

Chase, DZ, Chase, AF 1994. *Studies in the archaeology of Caracol, Belize* San Francisco, CA: Pre-Columbian Art Research Institute.

Chase, DZ, Chase, AF. 2017. Caracol, Belize, and Changing Perceptions of Ancient Maya Society. *Journal of Archaeological Research* **25:** 185-249.

Cornec, JH. 2003. Geology map of Belize, 1:750,000. Belmopan, Belize: Geology and Petroleum Office.

Cornuet, J-M, Pudlo, P, Veyssier, J, Dehne-Garcia, A, Gautier, M, Leblois, R, Marin, J-M, Estoup, A. 2014. DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* **30:** 1187-1189.

Cornuet, JM, Luikart, G. 1996. Description and Power Analysis of Two Tests for Detecting Recent Population Bottlenecks From Allele Frequency Data. *Genetics* **144:** 2001-2014.

Do, C, Waples, RS, Peel, D, Macbeth, GM, Tillett, BJ, Ovenden, JR. 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources* **14:** 209-214.

Dorsey, BL, Gregory, TJ, Sass, C, Specht, CD. 2018. Pleistocene diversification in an ancient lineage: a role for glacial cycles in the evolutionary history of Dioon Lindl. (Zamiaceae). *American Journal of Botany* **105:** 1512-1530.

Evanno, G, Regnaut, S, Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14:** 2611-2620.

Excoffier, L, Lischer, HEL. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10:** 564-567.

Fisher, JD, King Jr, D 2015. *Stratigraphy of the Toledo Formation, Belize Basin, Southern Belize*.

Garza, JC, Williamson, EG. 2001. Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology* **10:** 305-318.

Graham, A 1998. *Studies in Neotropical Paleobotany. XI. Late Tertiary Vegetation and Environments of Southeastern Guatemala: Palynofloras from the Mio-Pliocene Padre Miguel Group and the Pliocene Herreria Formation*.

Griffith, MP, Calonje, M, Meerow, AW, Francisco-Ortega, J, Knowles, L, Aguilar, R, Tut, F, Sánchez, V, Meyer, A, Noblick, LR, Magellan, TM. 2017. Will the same ex situ protocols give similar results for closely related species? *Biodiversity and Conservation* **26:** 2951-2966.

Griffith, MP, Calonje, M, Meerow, AW, Tut, F, Kramer, AT, Hird, A, Magellan, TM, Husby, CE. 2015. Can a Botanic Garden Cycad Collection Capture the Genetic Diversity in a Wild Population? *International Journal of Plant Sciences* **176:** 1-10.

Hall, JA, Walter, GH. 2013. Seed dispersal of the Australian cycad Macrozamia miquelii (Zamiaceae): Are cycads megafauna-dispersed "grove forming" plants? *American Journal of Botany* **100:** 1127-1136.

Hardin, JW. 1954. The vegetation pattern around a Florida sinkhole. *Contributions to Science* **2:** 6-17.

Hoese, G. 2012. Sinkhole and karst habitat of Ceratozamia latifolia and related Mexican cycads. *Mem. New York Bot. Gard.* **106:** 82-92.

Ishihara-Brito, R, Awe, JJ, Chase, AF. 2011. Ancient Maya Cave Use at Caracol, Belize. *Mexicon* **33:** 151-158.

Kalinowski, ST. 2004. Counting Alleles with Rarefaction: Private Alleles and Hierarchical Sampling Designs. *Conservation Genetics* **5:** 539-543.

Kalinowski, ST. 2005. Hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes* **5:** 187-189.

Kay, JS, A. A.;Murphy, V.;Nghiem-Phu, L.;Calonje, M.;Griffith, M. P. 2011. Palma Corcho: A Case Study in Botanic Garden Conservation Horticulture and Economics. *Horttechnology* **21:** 474-481.

Kivela, M, Arnaud-Haond, S, Saramaki, J. 2015. EDENetworks: a user-friendly software to build and analyse networks in biogeography, ecology and population genetics. *Mol Ecol Resour* **15:** 117-122.

Kopelman, NM, Mayzel, J, Jakobsson, M, Rosenberg, NA, Mayrose, I. 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* **15:** 1179-1191.

Langella, O. 2010. Populations 1.2.31 (http://bioinformatics.org/populations/) Accessed 17 January 2012.

Larmon, JT, McDonald, HG, Ambrose, S, DeSantis, LRG, Lucero, LJ. 2019. A year in the life of a giant ground sloth during the Last Glacial Maximum in Belize. *Science advances* **5:** eaau1200-eaau1200.

Lewis, PO, Zaykin, D. 2001. Genetic Data Analysis (GDA): Computer program for the analysis of allelic data, version 1.1. Free program distributed by the authors online at http://phylogeny.uconn.edu/software/.

Li, YL, Liu, JX. 2018. StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Mol Ecol Resour* **18:** 176-177.

Maruyama, T, Fuerst, PA. 1985. Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. *Genetics* **111:** 675-689.

Medina-Villarreal, A, González-Astorga, J, Espinosa de los Monteros, A. 2019. Evolution of Ceratozamia cycads: a proximate-ultimate approach. *Molecular Phylogenetics and Evolution***:** 106530.

Meerman, J, Sabido, W 2001. *Central American Ecosystems Map: Belize. Volume 2* Belize City, Belize: Programme for Belize.

Meerow, AW, Francisco-Ortega, J, Calonje, M, Griffith, MP, Ayala-Silva, T, Stevenson, DW, Nakamura, K. 2012. *Zamia* (Cycadales: Zamiaceae) on Puerto Rico: Asymmetric genetic differentiation and the hypothesis of multiple introductions. *American Journal of Botany* **99:** 1828-1839.

Meerow, AW, Nakamura, K. 2007. Ten microsatellite loci from Zamia integrifolia (Zamiaceae). *Molecular Ecology Notes* **7:** 824-826.

Meerow, AW, Salas-Leiva, DE, Calonje, M, Francisco-Ortega, J, Griffith, MP, Nakamura, K, Jiménez-Rodríguez, F, Lawrus, J, Oberli, A. 2018. Contrasting demographic history and population structure of *Zamia* (Cycadales: Zamiaceae) on six islands of the Greater Antilles suggests a model for population diversification in the Caribbean Clade of the genus. *International Journal of Plant Sciences* **0:** 000-000.

Miller, T. 1990. Caves and caving in Belize: an overview. *Caving* **46:** 2-4.

Miller, T 1996. *Geologic and hydrologic controls on karst and cave development in Belize*.

Nei, M, Tajima, F, Tateno, Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19:** 153-170.

O'Brien, CW, Tang, W. 2015. Revision of the New World cycad weevils of the subtribe Allocorynina, with description of two new genera and three new subgenera (Coleoptera: Belidae: Oxycoryninae). *Zootaxa* **3970:** 1-87.

Paetkau, D, Calvert, W, Stirling, I, Strobeck, C. 1995. Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology* **4:** 347-354.

Paetkau, D, Slade, R, Burden, M, Estoup, A. 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology* **13:** 55-65.

Pascual, M, Chapuis, MP, Mestres, F, Balanya, J, Huey, RB, Gilchrist, GW, Serra, L, Estoup, A. 2007. Introduction history of *Drosophila subobscura* in the New World: a microsatellite-based survey using ABC methods. *Molecular Ecology* **16:** 3069-3083.

Peakall, R, Smouse, PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* **28:** 2537-2539.

Peakall, ROD, Smouse, PE. 2006. genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6:** 288-295.

Pina-Martins, F, Silva, DN, Fino, J, Paulo, OS. 2017. Structure_threader: An improved method for automation and parallelization of programs structure, fastStructure and MavericK on multicore CPU systems. *Molecular Ecology Resources* **17:** e268-e274.

Piry, S, Alapetite, A, Cornuet, J-M, Paetkau, D, Baudouin, L, Estoup, A. 2004. GENECLASS2: A Software for Genetic Assignment and First-Generation Migrant Detection. *Journal of Heredity* **95:** 536-539.

Piry, S, Luikart, G, Cornuet, J-M. 1999. Computer note. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. *Journal of Heredity* **90:** 502-503.

Polk, JS, van Beynen, PE, Reeder, PP. 2007. Late Holocene environmental reconstruction using cave sediments from Belize. *Quaternary Research* **68:** 53-63.

Pritchard, JK, Stephens, M, Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155:** 945-959.

Pritchard, JKS, M.;Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155:** 945-959.

Puechmaille, SJ. 2016. The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Molecular Ecology Resources* **16:** 608-627.

Rambaut, A, Suchard, MA, Xie, D, Drummond, AJ. 2014. Tracer v1. 6. *Available from* http://tree.bio.ed.ac.uk/software/tracer/.

Rannala, B, Mountain, JL. 1997. Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences* **94:** 9197-9201.

Rousset, F. 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources* **8:** 103-106.

Rousset, F, Raymond, M. 1995. Testing heterozygote excess and deficiency. *Genetics* **140:** 1413-1419.

Salas-Leiva, DE, Meerow, AW, Calonje, M, Francisco-Ortega, J, Griffith, MP, Nakamura, K, Sanchez, V, Knowles, L, Knowles, D. 2017. Shifting Quaternary migration patterns in the Bahamian archipelago: Evidence from the *Zamia pumila* complex at the northern limits of the Caribbean island biodiversity hotspot. *American Journal of Botany* **104:** 757-771.

Sauro, U. 2012. Closed depressions in karst areas. In: White, WB, Culver, DC eds. *Encyclopedia of Caves.* Second Edition ed: Academic Press.

Slatkin, M. 2008. Linkage disequilibrium — understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics* **9:** 477.

Slatkin, M, Excoffier, L. 1996. Testing for linkage disequilibrium in genotypic data using the Expectation-Maximization algorithm. *Heredity (Edinb)* **76 (Pt 4):** 377-383.

Stevenson, DW. 2010. *Zamia prasina* Accessed 17 April 2019.

Tang, W. 1989. Seed dispersal in the cycad *Zamia pumila* in Florida. *Canadian Journal of Botany* **67:** 2066-2070.

Tang, W. 1993. Seed dispersal and reproductive effort in *Zamia pumila*. In: Stevenson, DW, Norstog, K eds. *The biology, structure, and systematics of the cycadales. Proceedings of Cycad 90, the second international conference on cycad biology.* Milton, Queensland: Palm & Cycad Societies of Australia Ltd.

Tang, W, Xu, G, O'Brien, CW, Calonje, M, Franz, NM, Johnston, MA, Taylor, A, Vovides, AP, Pérez-Farrera, MA, Salas-Morales, SH, Lazcano-Lara, JC, Skelley, P, Lopez-Gallego, C, Lindström, A, Rich, S. 2018. Molecular and Morphological Phylogenetic Analyses of New World Cycad Beetles: What They Reveal about Cycad Evolution in the New World. *Diversity* **10:** 38.

Terry, I, Tang, W, Taylor, AS, Donaldson, JS, Singh, R, Vovides, AP, Cibrián-Jaramillo, A. 2012. An Overview of Cycad Pollination Studies *Mem. New York Bot. Gard.* **106:** 352-394.

Toledo, VM. 1982. Pleistocene changes of vegetation in tropical Mexico. In: Prance, GT ed *Biological diversification in the tropics.* New York: Columbia University Press.

Van Oosterhout, C, Hutchinson, WF, Wills, DPM, Shipley, P. 2004. Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4:** 535-538.

Vovides, AP, Pérez-Farrera, MA, Iglesias, C. 2010. Cycad propagation by rural nurseries in Mexico as an alternative conservation strategy: 20 years on. *Kew Bulletin* **65:** 603-611.

Walker, JD, Geissman, JW, Bowring, SA, Babcock, LE. 2018. Geologic Time Scale v. 5.0 Accessed.

Wang, J. 2007. Triadic IBD coefficients and applications to estimating pairwise relatedness. *Genetical Research* **89:** 135-153.

Wang, J. 2011. Coancestry: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources* **11:** 141-145.

Waples, RS, Do, C. 2008. ldne: a program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources* **8:** 753-756.

Williamson-Natesan, EG. 2005. Comparison of methods for detecting bottlenecks from microsatellite loci. *Conservation Genetics* **6:** 551-562.

Wilson, GA, Rannala, B. 2003. Bayesian Inference of Recent Migration Rates Using Multilocus Genotypes. *Genetics* **163:** 1177-1191.

Yang, SL, Meerow, AW. 1996. The *Cycas pectinata* (Cycadaceae) complex: Genetic structure and gene flow. *International Journal of Plant Sciences* **157:** 468-483.

Table 1. Characteristics of sampled populations of *Zamia decumbens*

Population ID	**# of individuals sampled**	**Herbarium vouchers**	**Locality and notes**	**Habitat type (Meerman & Sabido, 2001)**
Cave	52	Calonje et al. BZ14-003 (FTG)	Belize: Toledo District: Maya Mountains. Columbia Forest Reserve, Outside cave entrance ca. 2.14 km SSW of Edwards Central.	Tropical evergreen broad-leaved submontane forest on rolling karstic hills
Hilltop	30	Calonje & Griffith BZ14-050 (FTG)	Belize: Toledo District: Maya Mountains. Hills South of Pueblo Viejo.	Tropical evergreen broadleaf lowland forest over steep calcareous hills
Sinkhole 1	178	Calonje et al. BZ08-201 (FTG)	Belize: Toledo District: Maya Mountains. Inside sinkhole N of San Jose (Hawaii) Village, 350-400 m elevation	Tropical evergreen broad-leaved lowland hill forest, *Callophyllum* variant
Sinkhole 2	179	Calonje et al. BZ08-222 (FTG)	Belize: Toledo District: Maya Mountains. Columbia Forest Reserve, Inside sinkhole N of Edwards Central Camp. 700 m elevation.	Tropical evergreen broad-leaved submontane forest on rolling karstic hills

Table 2. Descriptive statistics of Simple Sequence Repeat-based genetic variation of 10 SSR loci across 4 populations of *Zamia decumbens*

Locus	***N***		***N_a***		***N_e***			
	Mean	**SE**	**Mean**	**SE**	**Mean**	**SE**	***F_{IS}***	***F_{ST}***
Zam28	109.750	39.947	3.500	0.289	2.107	0.303	0.015	0.205
Zam33	109.750	39.947	3.250	0.479	1.964	0.312	-0.026	0.311
Zam53	109.750	39.947	15.500	2.102	5.914	1.837	-0.005	0.142
Zam59	108.500	40.736	5.750	1.109	2.882	0.275	0.015	0.112
Zam60	109.750	39.947	5.000	0.000	3.160	0.311	0.038	0.117
Zam61	109.000	39.803	3.250	0.250	1.840	0.225	-0.038	0.364
Zfg23	109.750	39.947	7.500	1.936	2.927	0.341	0.010	0.107
Zfg25	109.750	39.947	6.500	1.190	3.723	0.635	-0.027	0.119
Zfg32	109.750	39.947	2.750	0.479	1.948	0.219	0.027	0.095
Zfg33	109.000	39.803	8.500	1.555	4.402	0.340	-0.002	0.079
All loci	109.525	11.103	6.150	0.667	3.087	0.271	0.001	0.165
SE							0.008	0.031

Note: N = Effective sample size for the analysis, Na = number of alleles per locus, Ne = number of effective alleles, F_{IS} = Inbreeding coefficient, F_{ST} = coancestry coefficient

Table 3. Descriptive statistics of Simple Sequence Repeat-based genetic variation across four populations of *Zamia decumbens*

Population name	N_a	**AR**	**PA (freq. > .25)**	**PAR**	H_o	H_e	f	**Wilcoxon test**	M	**LD**	r
Cave (n=52)	7.4 ± 1.6	4.09	1.1 ± 0.38 (0)	0.95	0.72 ± 0.04	0.70 ± 0.04	0.11	E (P <0.001)	1.00	11	0.195
Hilltop (n=30)	6.0 ± 1.4	3.28	0.5 ± 0.31 (0)	0.62	0.53 ± 0.07	0.54 ± 0.07	0.31	n.s.	1.00	5	0.416
Sinkhole1 n=178)	4.0 ± 0.7	2.89	0.2 ± 0.13 (0)	0.55	0.58 ± 0.05	0.58 ± 0.05	0.20	E (P <0 .001)	0.70	19	0.355
Sinkhole2 (n=179)	7.2 ± 1.3	3.47	1.4 ± 0.54 (0)	0.69	0.60 ± 0.05	0.62 ± 0.05	0.18	E (P < 0.05)	1.00	28	0.250

Note. N_a = mean number of alleles across all loci; AR = mean allelic richness corrected for differences in sample size; PA = mean number of private alleles across all loci; PAR = mean private allelic richness corrected for differences in sample size; H_o = observed heterozygosity; H_e = expected heterozygosity; f = inbreeding coefficient; Wilcoxon test = one tailed Wilcoxon signed-rank test for heterozygosity excess (detects recent bottlenecks), n.s. = not significant; M = Garza and Williamson (2001) coefficient; LD = linkage disequilibrium, reported as the number of paired loci per population testing positively for LD at $P < 0.001$; r = mean pairwise relatedness (Wang, 2007).

Table 4. Permuted analysis of molecular variance results with simple sequence repeats (SSRs) across four *Zamia decumbens* populations

	df	**Sum of squares**	**Variance components**	**Percentage**
Among populations	3	463.150	0.795	21%
Within populations	874	2660.188	3.044	79%
Total	877	3123.338	3.838	100%

Table 5. Pairwise permuted FST values between populations of *Zamia decumbens*

	Cave	**Hilltop**	**Sinkhole1**	**Sinkhole2**
Cave	—	—	—	—
Hilltop	0.103	—	—	—
Sinkhole1	0.204	0.268	—	—
Sinkhole2	0.116	0.241	0.223	—

Table 6. Estimates of effective population size (N_e), number of generations since population coalescence (t), mutation rate per generation (µmic), the shape parameter of the gamma distribution describing the heterogeneity of rates among individual loci (pmic), and the single insertion nucleotide rate (snmic), estimated from DIYABC analyses of simple sequence repeat data for *Zamia decumbens*

Parameter	Mean	Median	Mode
N_e (Cave)	34.72	34.60	32.11
N_e (Hilltop)	240.90	233.30	222.30
N_e (Sinkhole1)	41.96	41.67	41.59
N_e (Sinkhole2)	47.16	47.00	46.62
t1	2353	2182	1810
t2	6067	6144	6758
t3	6117	6198	6794
µmic_1	0.000301	0.000297	0.000294
pmic_1	0.185300	0.181300	0.167400
snimic_1	1.02E-07	8.58E-08	5.06E-08

Table 7. Estimates (means) of effective population sizes (Ne) and years since coalescence of *Zamia decumbens* populations based on the most optimal scenario of divergence (Fig. 5) with DIYABC using SSR data. The range of years since coalescence estimated using generation times of 15 and 30 years

	N_e	Coalescence node	Years
Sinkhole2	47.5	t3	95,565–191,130
Sinkhole1	47.16	t2	92,760–185,520
Cave	34.72	t1	38,895–77,790
Hilltop	240.90	…	…

Table 8. Effective population sizes (Ne) of *Zamia decumbens* populations as calculated with Ne Estimator based on the LD method. Estimated Ne values are provided based on three different thresholds for the lowest allele frequency used, the 95% confidence intervals in parentheses are based on jackknifing on loci; n: sample size; "infinite" values, usually resulting from sampling error (Do et al., 2014), represent cases where there is no evidence that variation in the genetic characteristics are caused by a finite number of parental individuals

Population	*n*	Effective population size (N_e)		
		Frequency threshold:		
		0.05	0.02	0.01
Cave	52	22.6 (17.1–30.6)	38.5 (29.7–51.7)	35.1 (27.4–46.2)
Hilltop	30	86.8 (26.2–infinite)	163.2 (38.6–infinite)	69.3 (33.9–502.8)
Sinkhole1	178	16.1 (2.9–47.1)	21.7 (5.5–57.9)	23.7 (6.4–65.3)
Sinkhole2	179	32.3 (17.8–58.1)	39.7 (25.0–63.5)	47.5 (31.3–73.7)

Table 9. Bayesian assessment of migration among and within *Z. decumbens* populations. The numbers are the mean proportions of individuals from each source

Migration direction	**Mean**	**95% HPD Interval**
Cave→Cave	0.9665	[0.9326, 0.9944]
Cave→Hilltop	0.0162	[8.1685E-8, 0.0422]
Cave→Sinkhole1	0.0065	[1.8168E-6, 0.0194]
Cave→Sinkhole2	0.0108	[3.6082E-6, 0.03]
Hilltop→Cave	0.0164	[1.1044E-6, 0.0449]
Hilltop→Hilltop	0.9621	[0.9236, 0.9957]
Hilltop→Sinkhole1	0.0100	[3.4668E-7, 0.0294]
Hilltop→Sinkhole2	0.0115	[3.243E-7, 0.0335]
Sinkhole1→Cave	0.0020	[1.6165E-7, 5.9199E-3]
Sinkhole1→Hilltop	0.0021	[3.7431E-7, 6.1755E-3]
Sinkhole1→Sinkhole1	0.9940	[0.9875, 0.9993]
Sinkhole1→Sinkhole2	0.0020	[2.1301E-7, 5.7728E-3]
Sinkhole2→Cave	0.0051	[4.1923E-7, 0.0127]
Sinkhole2→Hilltop	0.0059	[1.1484E-4, 0.0131]
Sinkhole2→Sinkhole1	0.0021	[1.6076E-7, 6.3526E-3]
Sinkhole2→Sinkhole2	0.9869	[0.9755, 0.9966]

Table 10. First-generation migrants identified among four populations of *Zamia decumbens* in Geneclass2 analysis

Individual	**Sample location**	**Origin of ancestry**	**Log(L) of ancestry**
SH2-013	Sinkhole2	Cave	12.504
SH2-124	Sinkhole2	Cave	10.301
SH2-135	Sinkhole2	Cave	10.379
SH2-140	Sinkhole2	Cave	12.771
Cave-10	Cave	Hilltop	11.655
SH2-137	Sinkhole2	Cave	14.151
Hillltop-26	Hilltop	Cave	9.746
SH2-121	Sinkhole2	Cave	13.72
Hilltop-05	Hilltop	Cave	11.345
SH2-148	Sinkhole2	Cave	11.843

Fig. 1. Geographic distribution of the four populations of *Zamia decumbens* studied in Toledo District of Belize. In the inset map, the red rectangle indicates the study area displayed in the larger map, the black dashed line represents the entire extent of occurrence for the species. SH = Sinkhole

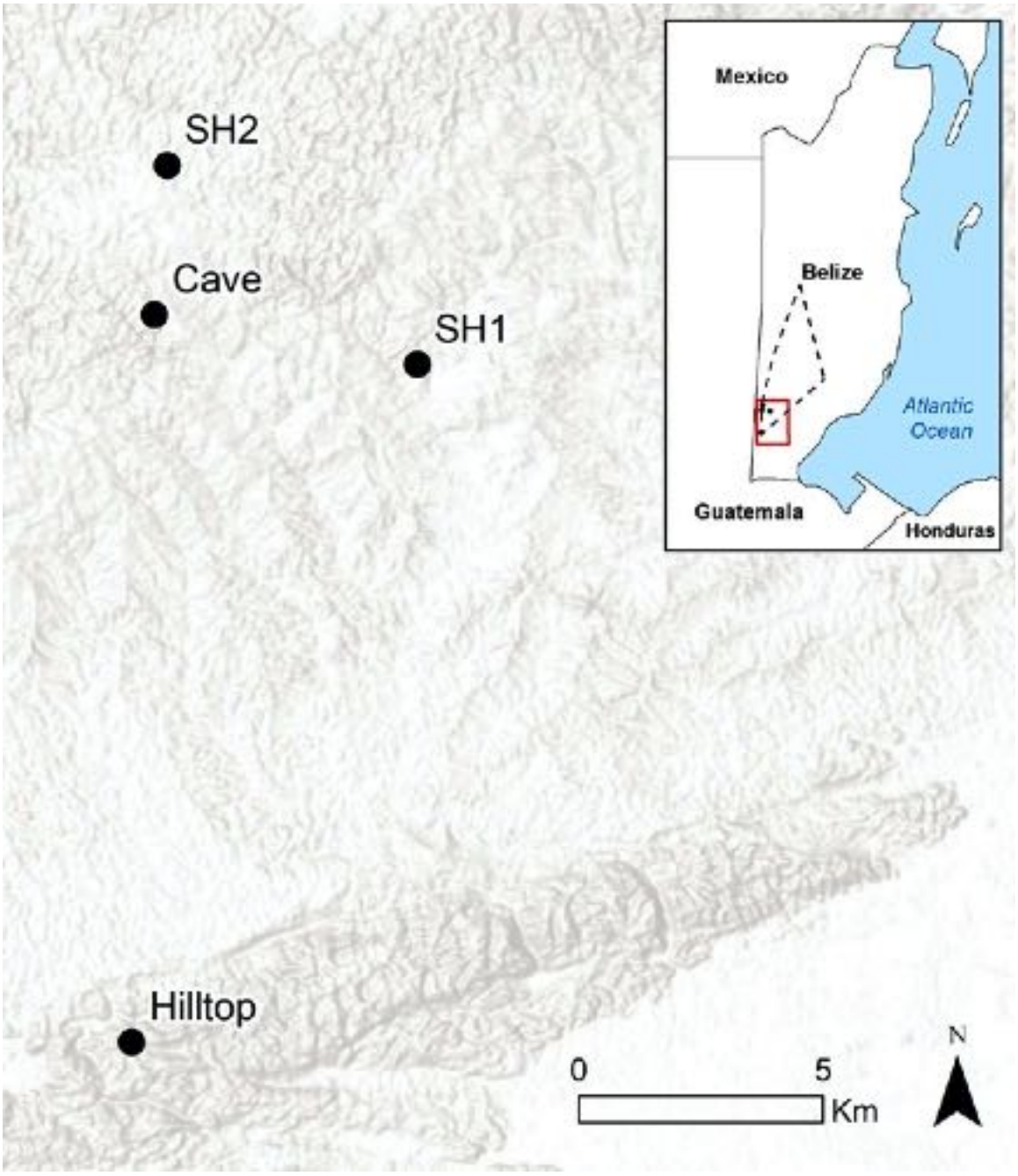


Fig. 2. Populations of *Zamia decumbens* sampled in this study. A & B: Sinkhole1, C&D: Sinkhole2, E&F: Cave, G: Hilltop.

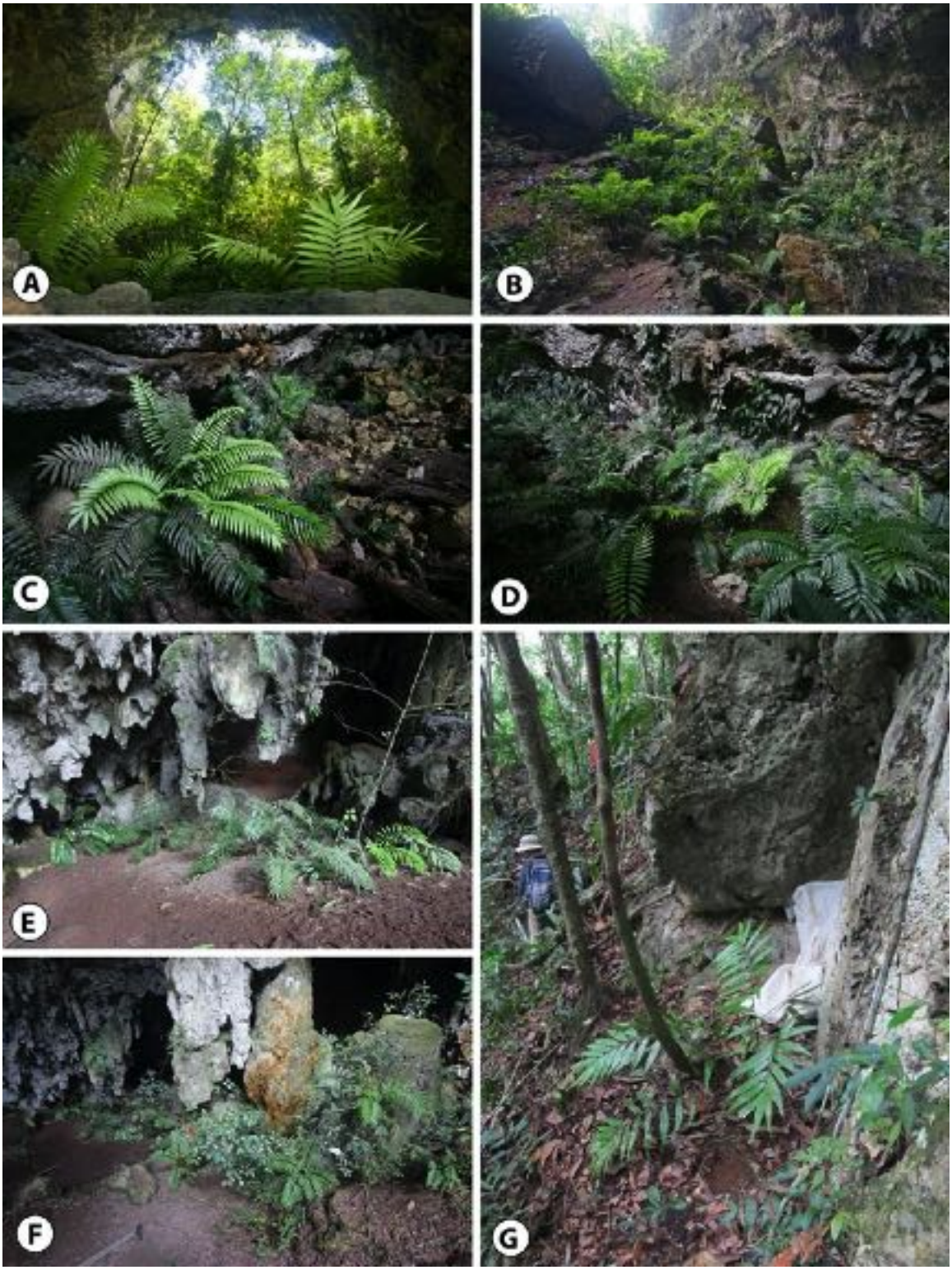


Fig. 3. STRUCTURE cluster histograms of four populations of *Zamia decumbens* for K=2, the optimum number of distinct genetic populations as resolved with Δ*K* statistics k estimators; and K=3 and K=4 as resolved with MMK statistics.

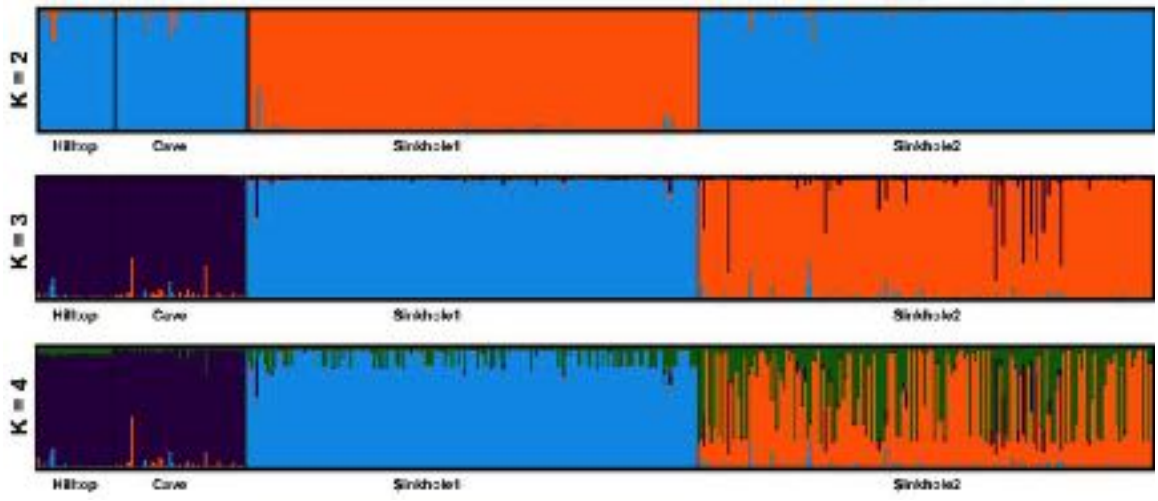


Fig. 4. Results from DIYABC. (A) Optimal scenario of ten scenarios (See Figs. A1 & A2) analyzed with DIYABC. This invasion scenario shows that the two sinkhole populations studied, by far the largest and most robust, were the first established, followed by CAVE and Hilltop populations; Scenario comparisons with (B) logistic regression and (C) direct approach.

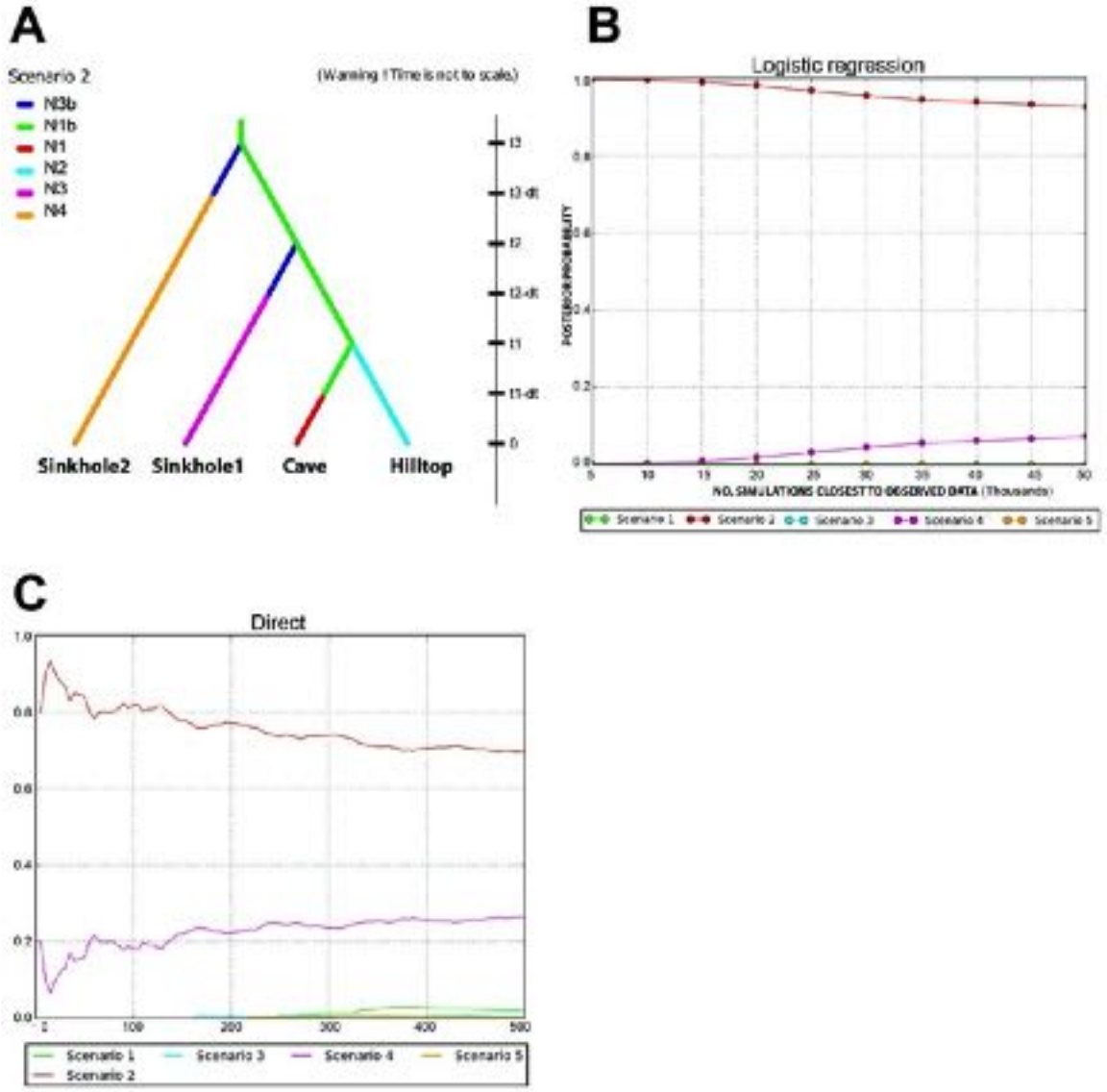


Fig. 5. Automatically thresholded network of four *Zamia decumbens* populations based on simple sequence repeat F_{ST} generated with the software EDENetworks (Kivelä et al. 2015). The layout of the network nodes is based on the true geographic coordinates of the populations and is visualized over a topographic map. Network node size corresponds to betweenness centrality, and edge weight is inversely proportional to pairwise F_{ST} values, with thicker edges indicating a closer relationship between populations.

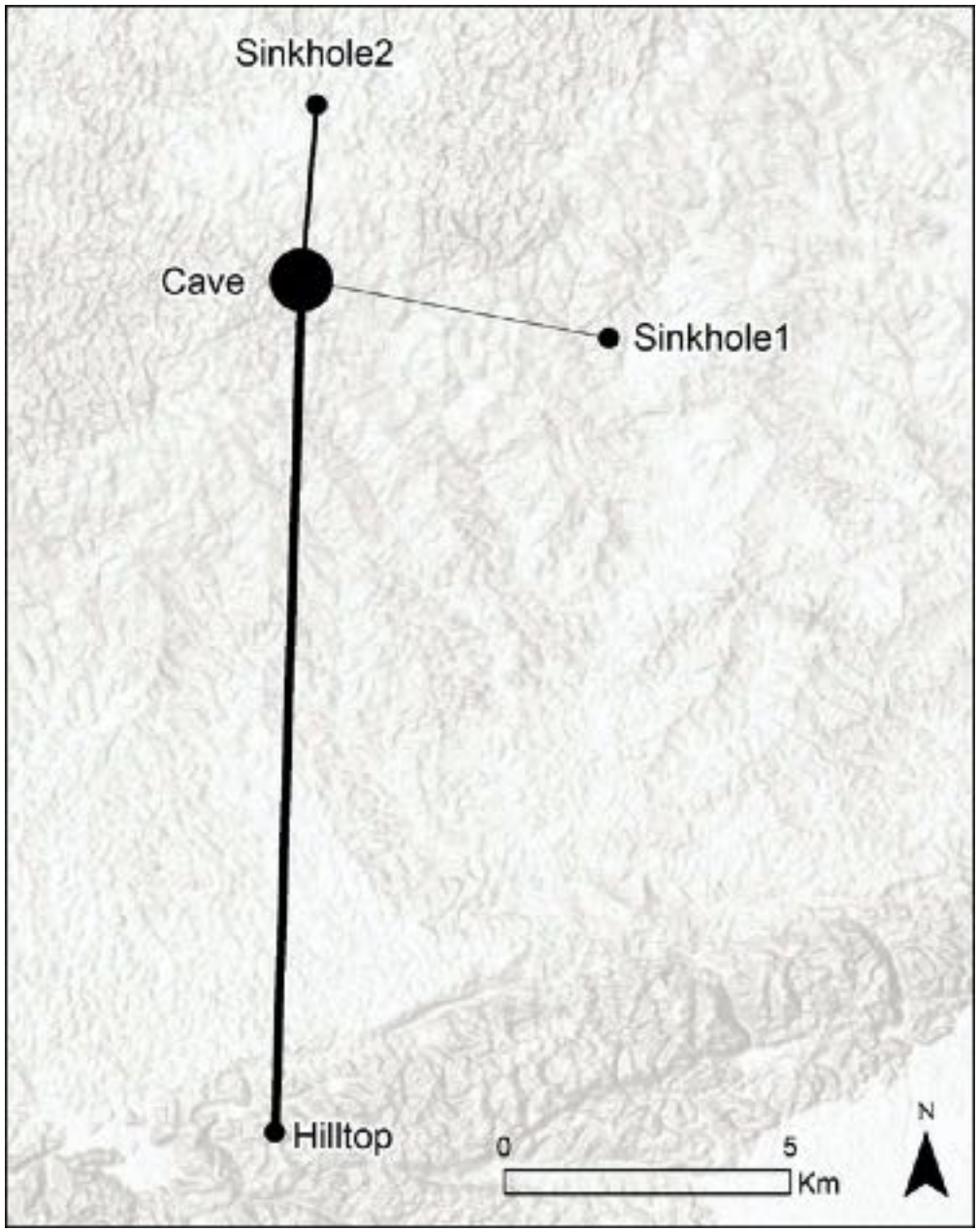
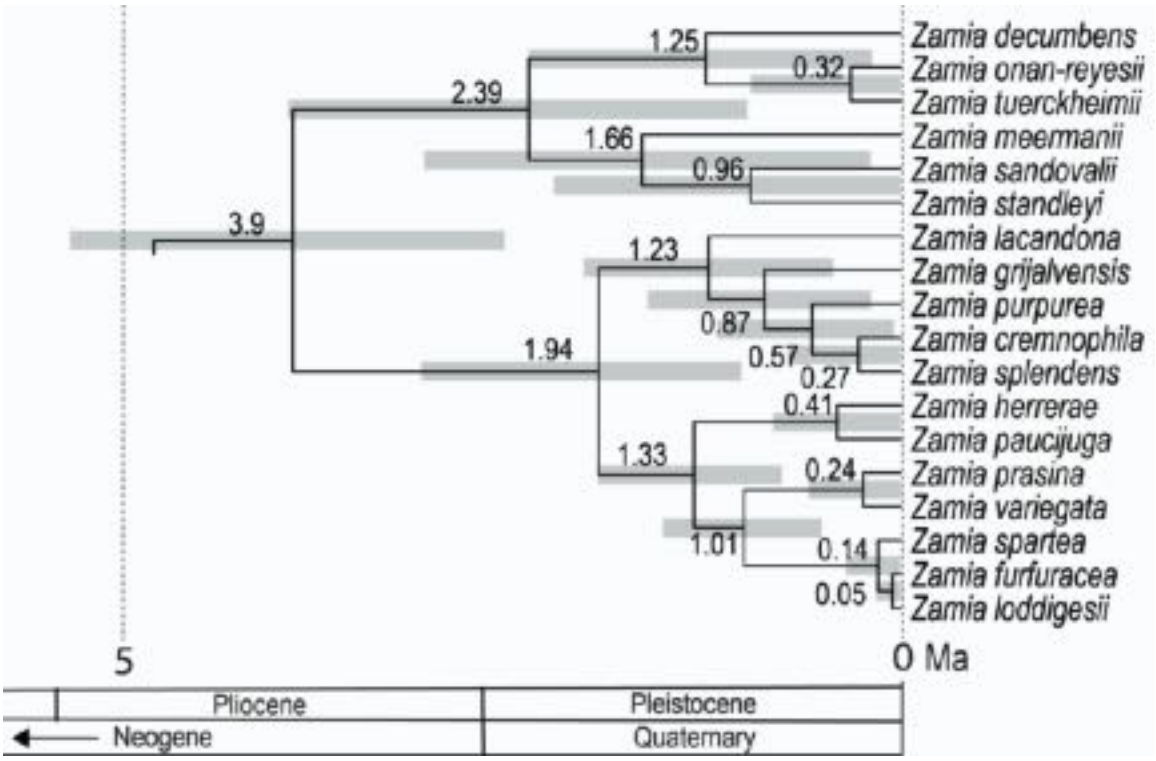


Fig. 6. Phylogenetic position of *Zamia decumbens* within the Tuerckheimii clade from a 10 locus (9 single-copy nuclear genes and 1 chloroplast) time-calibrated species-tree of the genus *Zamia* (Calonje et al., 2019).



APPENDIX A: SUPPLEMENTARY FIGURES

Fig. A1. Population divergence invasion scenarios tested using direct approach analysis of simple sequence repeats with DIYABC (Cornuet et al. 2014). Scenario 2 was the most probable among the 10 divergence scenarios tested.

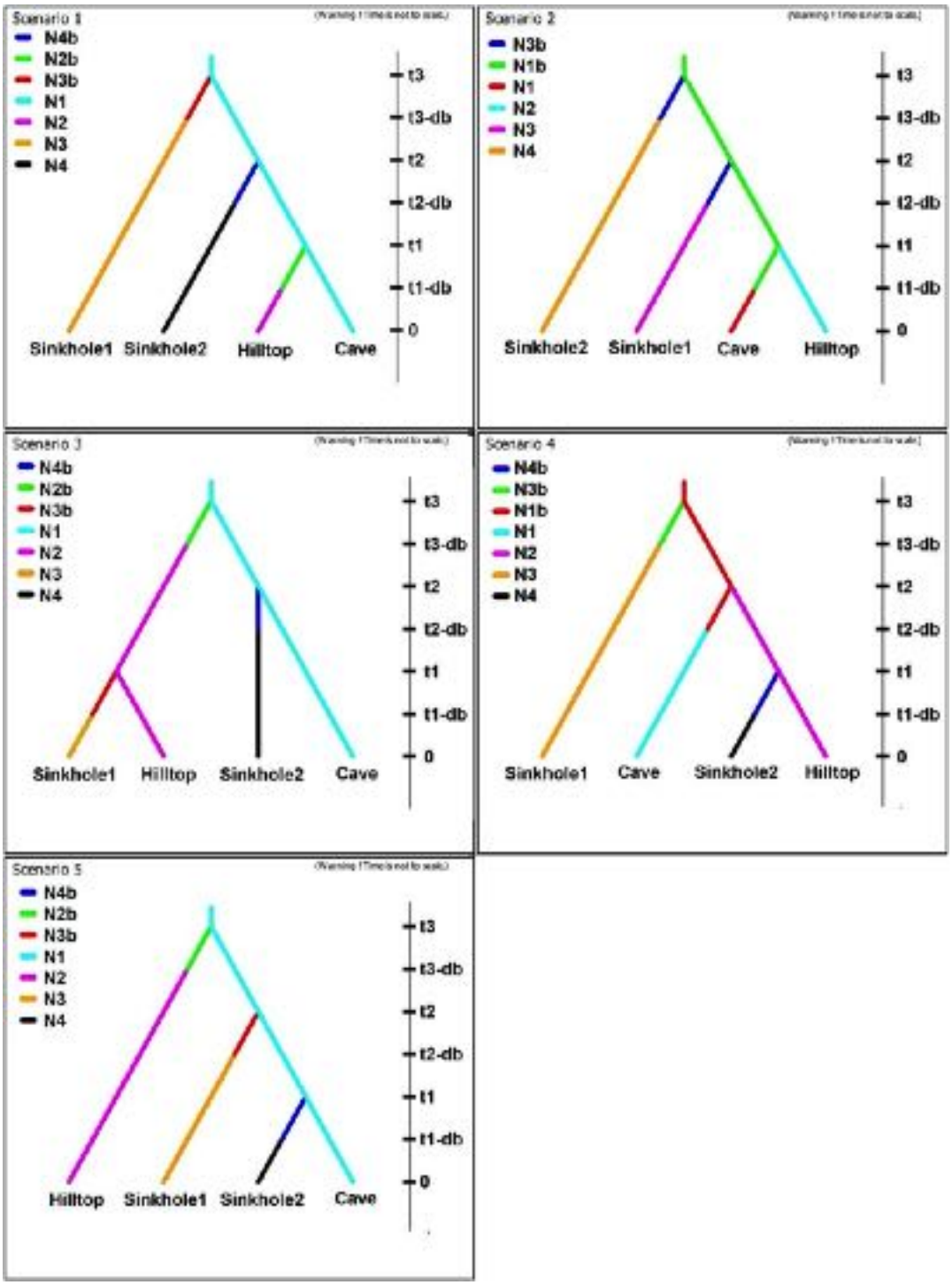
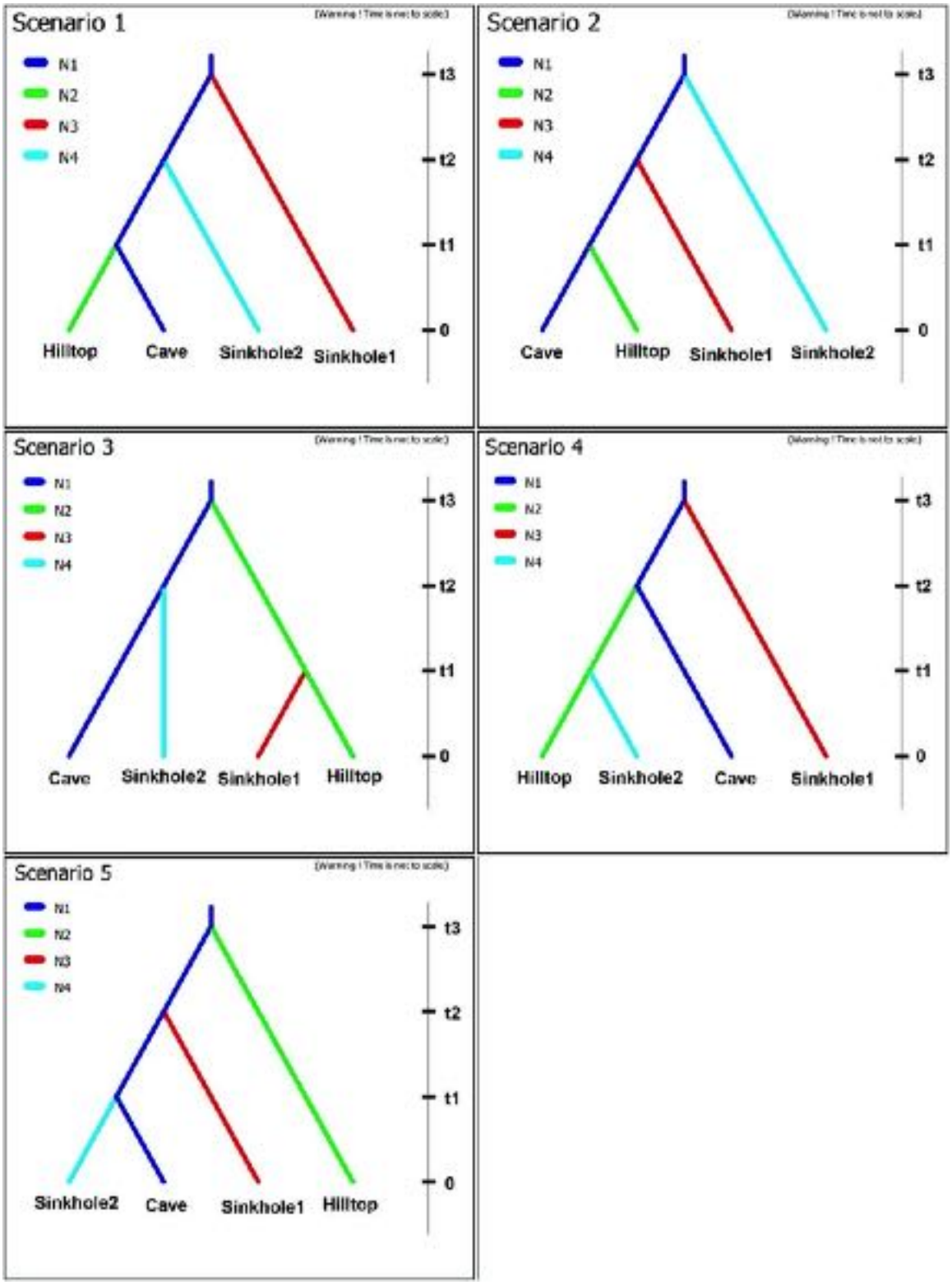


Fig. A2. Population divergence vicariance scenarios tested using direct approach analysis of simple sequence repeats with DIYABC (Cornuet et al. 2014).



CHAPTER III

Cycad biodiversity in the Bahamas Archipelago and conservation genetics of the threatened *Zamia lucayana* (Zamiaceae)

MICHAEL CALONJE BAZAR[a,b*], ALAN W. MEEROW[c], LINDY KNOWLES[c], DAVID KNOWLES[d] and JAVIER FRANCISCO-ORTEGA[a, e]

[a]Department of Biological Sciences, Florida International University, Miami, FL 33199, USA; [b]Montgomery Botanical Center, Coral Gables, FL 33156, USA; [c]USDA-ARS-SHRS, Miami, FL 33158, USA; [d]Bahamas National Trust, Nassau, The Bahamas; [e]Bahamas National Trust, Abaco National Park, Abaco, The Bahamas; [e]Fairchild Tropical Botanic Garden, Coral Gables, FL 33156, USA.

[*] For correspondence: michaelc@montgomerybotanical.org

Chapter 3 was published in Oryx 47(2): 190–198 ; doi:10.1017/S0030605312000129, with the same title and authors.

ABSTRACT

A conservation assessment for the three cycad species native to the Bahamas Islands is presented. Results are based on field surveys on all islands where these species occur. *Zamia angustifolia* is native to Eleuthera, *Zamia integrifolia* is native to Abaco, Andros, Eleuthera, Grand Bahama and New Providence, and *Z. lucayana* is endemic to Long Island. *Zamia angustifolia* is of the highest conservation concern because of the small number of adult plants, its restricted distribution and the extensive development occurring within its habitat. *Zamia integrifolia* also has a restricted distribution on Eleuthera and Grand Bahama and, although threatened by urban development in New Providence, it is relatively common on Abaco and Andros. *Zamia lucayana* comprises three populations within a narrow strip of land of c. 1 km^2; we propose a reassignment of its current conservation status from Endangered to Critically Endangered. We assessed the genetic structure of *Z. lucayana* based on 15 polymorphic microsatellite DNA loci; this indicated that the three known populations should be considered a single management unit. However, the high number of private alleles suggests that genetic drift, indicative of recent fragmentation, is progressing. We propose in situ conservation strategies, and we also collected germplasm from a total of 24 populations of these three cycad species, for ex situ conservation.

INTRODUCTION

There are c. 331 species of cycads in 10 genera (Osborne *et al.*, 2012) and they are often referred to as living fossils because they first appeared in the Triassic (Norstog & Nichols, 1997). These gymnosperms are rapidly disappearing because of habitat loss and their popularity in the ornamental plant trade, which has led to drastic declines in population size (Osborne, 1995). Most cycad species are currently on the IUCN Red List (IUCN, 2011) and their trade is regulated by CITES (CITES, 2011). All cycads are dioecious and their pollination and dispersal biology constrain both gene flow and population recruitment to relatively small geographical areas (Norstog & Nichols, 1997).

There are two cycad genera in the Caribbean island biodiversity hotspot: *Microcycas* and *Zamia*. The former is a monotypic Critically Endangered genus restricted to Cuba (Pinares, 2009). *Zamia* (c. 71 species) has a wider distribution from Bolivia to Central America, Mexico, Florida and the West Indies. The Caribbean island species form a monophyletic group with c. nine recognized species, known as the *Zamia*

pumila complex (Stevenson, 1987). Three of these taxa occur in the Bahamas archipelago: *Z. angustifolia* (also present in Cuba), *Z. integrifolia* (also occurring in Cuba and Florida), and *Z. lucayana* (Fig 1). The latter (Fig 2) is endemic to Long Island. *Z. tenuis* is another species reported for the Bahamas but it is only known from a single specimen collected in the early 19th century from a plant cultivated at Berlin Botanical Garden. *Z. tenuis* was formerly considered a synonym of *Z. angustifolia* by Britton and Millspaugh (1920) and is currently treated as a synonym of *Z. integrifolia* (Osborne *et al.*, 2012). The Commonwealth of the Bahamas consists of 22 inhabited islands, c. 700 cays and nearly 2,400 small islets (Albury, 1975). This archipelago has a fast growing population and a strong tourism industry, which represent challenges to the protection of its unique environment (Anonymous, 2002). It is estimated that this archipelago has 110 endemic seed plant species (c. 9% of the native flora; Acevedo-Rodríguez & Strong, 2012). As far as we are aware no Bahamian endemic species has been the subject of a comprehensive study focusing on (1) its conservation status, (2) challenges for management, (3) conservation genetics, (4) ex situ conservation, and (5) current threats. Such studies are particularly relevant to the Bahamas as this archipelago is part of the Caribbean island biodiversity hotspot (Maunder *et al.*, 2008). Here, we formulate conservation assessments of Bahamian *Zamia* species, with special emphasis on the Long Island endemic *Z. lucayana*.

MATERIALS AND METHODS

COLLECTING SITES

Prior to our field studies we compiled a list of historical sites for the target species as recorded by collections deposited in seven herbaria (BW, BNH, C, FTG, GH, MO, NY; acronyms follow Thiers, 2011). In addition, we consulted floristic and taxonomic treatments that include information concerning the distribution of the genus in the Bahamas (Britton, 1907, Britton, 1909, Eckenwalder, 1980, Correll & Correll, 1982, Hill, 1974), and gathered provenance information from living collections at Montgomery Botanical Center and Fairchild Tropical Botanic Garden. Between December 2009 and July 2011 we conducted field studies on the five islands where the three species occur (Table 1, Fig 1). Each island has only one species, except Eleuthera, where *Z. angustifolia* and *Z. integrifolia* co-occur, although their distribution ranges do not

overlap (Fig 1). We visited a total of 24 populations (Fig 1, Table 1) and we also explored areas where the species were not previously reported.

CONSERVATION ASSESSMENTS AND DEMOGRAPHIC SURVEYS

For each site and island we gathered information pertinent to the cycads' conservation status and threats. Field studies were more exhaustive for *Z. lucayana*, for which we estimated total distribution area, number of adult and juvenile plants, and the approximate ratio of males to females. Juveniles are readily identified by their overall smaller size, absence of remains of reproductive structures, lower number of leaves (usually < 4) on the stem, and lower number (usually < 5 pairs) of leaflets on the leaves. These demographic data were obtained through a census of 10 × 10 m plots that covered c. 25% of the populations. Locations of individual plants occurring at the edges of the populations were determined with a global positioning system; these points were imported to *ArcMap v. 10* (ESRI, Redlands, USA) and used to determine area of occupancy and extent of occurrence (IUCN, 2011) for the species.

MICROSATELLITE DNA ANALYSIS

We used 15 DNA microsatellites to investigate the genetic structure of *Z. lucayana*. The three known populations of *Z. lucayana* were included (Table 2) and we sampled 33–46 individuals per population. DNA was isolated following the protocols described by Meerow and Nakamura (2007). Microsatellite loci were developed in two ways. A modified version of the enrichment/hybridization method of Edwards *et al.* (1996) was used with genomic DNA of Florida *Z. integrifolia* (Meerow & Nakamura, 2007; Meerow et al., 2007, 2012). The second methodological approach involved retrieving of microsatellite loci in the *Zamia* expressed sequence tag (EST) databases available in GenBank using the Simple-Sequence Repeat Identification Tool in *GRAMENE* (Ware *et al.*, 2002). Six of the 15 loci (Zam28, Zam33, Zam34, Zam39, Zam40, Zam45) have been previously published (Meerow & Nakamura, 2007). Protocols for PCR amplification and subsequent visualization of SSR fragments follow Meerow and Nakamura (2007).

Descriptive statistics (Table 2) were generated with *GenAlEx v. 6.41* (Peakall & Smouse, 2006). Tests for Hardy–Weinberg equilibrium and the U test (Rousset & Raymond, 1995) for heterozygote excess or deficiency were run with *GenePop v. 4.0* using 10,000 Monte Carlo Markov chain iterations (Guo &

Thompson, 1992). Linkage disequilibrium (LD) was tested for each population with *ARLEQUIN v. 3.5* (Excoffier & Lischer, 2010) using a likelihood-ratio test (Slatkin & Excoffier, 1996). A Monte Carlo Markov chain method was applied with 100,000 iterations, a burn-in of 10,000 and the significance level set at $P < 0.001$.

Analysis of molecular variance (AMOVA) among populations, and permuted calculation of F_{ST} (10,000 permutations) were generated with *GenAlEx.* An additional measure of gene diversity, D_{est} (Jost, 2008), was calculated with *SMOGD* (Crawford, 2010).

Genetic distance among populations and individuals was calculated with *POPULATIONS v. 1.2.31* (Langella, 2010), using *Da* (Nei, Tajima & Tateno, 1983). The population distance matrix was used for permuted (10,000 iterations) Mantel (1967) tests for isolation by distance following the methods of Smouse, Long and Sokal (1986) and Smouse and Long (1992), as implemented in *GenAlEx*. These distance coefficients were also used in principal coordinate analysis with *GenAlEx. BOTTLENECK 1.2.02* (Cornuet & Luikart, 1996) was used to test for recent genetic bottlenecks in the populations under both the infinite allele model (Kimura & Crow, 1964) and the two phase model (Di Rienzo *et al.*, 1994).

The Bayesian clustering programme *STRUCTURE v.2.3.3* (Pritchard, Stephens & Donnelly, 2000) was used to estimate the underlying genetic structure among populations. The *STRUCTURE* analyses were carried out on the University of Oslo Bioportal (Kumar *et al.*, 2009). *K* values of 1–15 were simulated across 20 replicate runs of 1,000,000 iterations after a burn-in of 100,000. The Δk method of Evanno, Regnaut and Goudet (2005) as implemented in *STRUCTURE HARVESTER* (Earl & vonHoldt, 2012) was used to determine the 'true' value of *K* across samples. After the likely level of *K* was estimated, a consensus Q-matrix from the 20 runs was constructed using *CLUMPP* ((Jakobsson & Rosenberg, 2007) for visualization with *DISTRUCT* (Rosenberg, 2004).

RESULTS

DISTRIBUTION PATTERNS

With few exceptions (Lubber's Quarters Cay off the coast of Abaco and in the vicinity of Eight Mile Rock, Grand Bahama), we found populations of *Zamia* in the areas recorded on the herbarium specimens and in the relevant literature. In our field surveys beyond the historical distribution ranges we did not find any new localities for the species, with the exception of those located on the cays near Abaco.

Zamia angustifolia is extremely rare and occurs only on coastal sand dunes of Eleuthera. We found only 150 plants in two sites within a small area of 0.34 km^2 near Gregory Town (Fig 1).

Zamia integrifolia is typically found in the understory of the Bahamian pine forests and occasionally in dry evergreen forests, except on Tilloo Cay and Eleuthera. In these two islands pine forests are not present and the species occurs on limestone bluffs in coastal thicket (Eleuthera) or in sandy coastal scrub (Tilloo Cay). The species is abundant in Abaco, relatively common in northern areas of North Andros and locally abundant in a few areas with unfragmented pine forest on New Providence. However, it is rare on Grand Bahama and Eleuthera (Fig 1).

Zamia lucayana occurs on the eastern coast of Long Island between the settlements of Hamilton's and Buckley's (Fig.). The three populations grow exclusively within a narrow strip (c. 6.5 km by 100 m wide) of coastal scrub vegetation on sandy soils in association with sea grape *Coccoloba uvifera*, occupying a total area of c. 1 km^2. The combined area of occupancy for the surveyed populations is 0.06 km^2 (Table 2). Approximately 80% of the fertile adult individuals surveyed were male and 20% female (Table 2). Approximately 27% of individuals were juvenile plants (Table 2). All female individuals with mature cones had nearly complete seed set, indicating efficient pollination. We observed hermit crabs *Coenobita clypeatus* feeding on the fleshy seed coats, perhaps accelerating germination and short-range dispersal (Fig 2). We estimated there are c. 980 adult individuals throughout the entire range of this species, with 240–400 plants per population (excluding the two outlier sites; Table 2).

POPULATION GENETICS OF *Z. LUCAYANA*

The 15 polymorphic microsatellite loci uniquely genotyped all 122 individuals assayed; i.e. there were no identical repeated multi-locus genotypes. The mean number of alleles was 4.3–5.3 (Table 2) but the populations had a relatively high number of private alleles, with seven for ZBLI3 and 15 for both ZBLI1 and ZBLI2 (Table 2). All populations are moderately heterozygous (Table 2). All but ZBLI1 are slightly inbred but variation in *F* was not significant. ZBLI1 also has the highest mean number of alleles per locus. Exact tests found no significant departure from Hardy–Weinberg equilibrium in any populations. Overall, there was 6% linkage disequilibrium among loci, the majority of which were concentrated in ZBLI2 (5%). ZBLI1 had no loci in LD; ZBLI3 had two.

Mean F_{ST} was highly significant (0.067, $P < 0.0001$) but indicates little differentiation among populations (Table 3). Similarly Jost's D_{est} estimate of differentiation is low, at 0.04–0.051 (Table 3). The mean number of migrants (Nm) determined by the F_{ST} method is 3.6 (range 3.07–3.94; Table 4), indicating gene flow among the populations. There is no significant isolation by distance among the populations. Over 90% of the genetic variation is within populations (Table 5). The Evanno method of determining the true *K* identified $K = 2$ as optimal across all *Z. lucayana* populations (Fig. 4). ZBL3 is located between ZBLI1 and ZBL2; it shows a closer genetic relationship to ZBLI1 but with some significant admixture from ZBLI2. Principal coordinate analysis indicates much the same as the Bayesian clustering (Fig. 5) and ZBLI3 mostly overlaps with ZBLI2. The first two coordinates of this analysis accounted for 43.44% of the variation.

Across all three *Zamia lucayana* populations there is only weak evidence of genetic bottlenecks. ZBLI1 and ZBLI2 only tested at $P < 0.05$ for the Sign test under the infinite allele model. ZPLI3 tested at $P < 0.05$ for both the Sign and Wilcoxon tests under the infinite allele model. With the two-phase model imposed there was no evidence of bottlenecks in any of the populations.

DISCUSSION

DISTRIBUTION PATTERNS AND CONSERVATION IMPLICATIONS

Zamia angustifolia is the species of highest conservation concern because of the small number of adult plants, its limited habitat extent, and the extensive housing development projects occurring in this area. The species also occurs in Cuba, where it is relatively abundant (González-Géigel, 2003). However, it is uncertain whether the narrow-leaved morph of Cuba is the same taxon that occurs on Eleuthera. The Bahamian populations are highly threatened and additional studies that include the Cuban populations are needed to determine the conservation status of this species.

Zamia lucayana was described by Britton (1907) based on a single individual reported in an unknown locality apparently near Clarence Town (Fig. 3). Later Hill (1974) rediscovered this species, finding a dense population in the vicinity of Hamilton's Settlement (Fig. 3). Our field studies confirmed the existence of this population and we discovered two additional populations near the Settlements of Buckley's and Petty's. We found that human activities have not yet had a major negative effect on the populations of *Z. lucayana*. The population at Hamilton's Settlement is within an area that has been exploited for sand mining but this activity is localized. All the locations of this species are accessible by road and are on prime ocean-front real estate that has been subdivided for sale, although few houses have been built.

New Providence and Grand Bahama are the two islands where *Zamia integrifolia* is at most risk. As urban development of Nassau expands westwards the remaining habitats of this species will be destroyed; its future on the island is thus uncertain. On Grand Bahama the species occurs near Freeport, at a few sites within an area with intense industrial development.

CONSERVATION GENETICS

Critically Endangered species usually have reduced numbers of populations and individuals per population (IUCN, 2011) and the genetic structure of their populations may therefore be severely influenced by genetic drift and inbreeding (Höglund, 2009). This results in a decrease in heterozygosity, with subsequent risk of inbreeding depression, and loss of potential adaptive alleles (Peterson & McCracken, 2005). Nevertheless, our data show that despite its Critically Endangered status (see below) the populations of *Zamia*

lucayana are still in Hardy–Weinberg equilibrium and are retaining heterozygosity, and are only moderately inbred and exhibit little differentiation.

Our results indicate that the three populations of *Zamia lucayana* on Long Island have been a single panmictic population throughout most of their history. The low levels of genetic differentiation (Table 3) and relatively high migration rate (Table 4) among the three populations indicate that gene flow has been historically high. However, the high number of private alleles (Table 2) indicates that genetic drift is occurring, especially in ZBLI1 and ZBLI2. Given the characteristically local dispersal of both pollen and seeds of *Zamia* (Norstog & Nichols, 1997) our expectations are that the populations will continue to fragment further if disturbed, and become more inbred. The genetic data indicate that the three populations of *Z. lucayana* should be considered as a single management unit. The population at Petty's has the highest levels of admixture, and therefore this population would be a good source of material for reintroducing the species to other areas of the island or for ex situ conservation.

Our data contrast with those recently reported for other Critically Endangered plant taxa from the Caribbean islands in which microsatellites have been used to determine genetic structure. Namoff *et al.* (2011) found strong evidence for genetic drift, inbreeding and moderate gene flow for the Critically Endangered palm *Pseudophoenix ekmanii*, endemic to the Dominican Republic. Geiger *et al.* (2014) reported that populations of *Ipomoea microdactyla* from highly fragmented and disturbed areas of South Florida exhibited significantly lower levels of genetic variation than those from the contiguous and well-preserved pine forest on Andros Island.

Walters and Decker Walters (1991) used isozyme data to determine levels of genetic variation within the *Zamia pumila* complex but found limited genetic variation with these markers. Our research suggests that microsatellites are more useful than isozymes to understand the population genetic structure of this species complex (Meerow & Nakamura, 2007; Meerow et al., 2007, 2012). This is supported by our results for *Z. lucayana*.

Most population genetic studies of cycads have been based on isozymes (reviewed by Pinares, 2009), with only three studies focusing on Critically Endangered species (*Dioon caputoi*, Cabrera-Toledo et al. 2008, 2010*; Microcycas calocoma*, Pinares et al., 2009). *Cycas debaoensis* is the only Critically Endangered cycad

(IUCN, 2011) for which conservation genetic studies based on microsatellites are available (Yang *et al.*, 2008).

The unusual patterns of genetic diversity detected in *Zamia lucayana* are also exhibited by these three other Critically Endangered species, which also exhibit relatively high levels of heterozygosity. These results appear to provide additional support for the hypothesis (Cabrera-Toledo, Gonzalez-Astorga & Vovides, 2008) that in cycads 'rarity is compatible with high levels of genetic diversity'. Unlike many seed plants, cycads are long lived and allogamous. These life-history characteristics may help reduce inbreeding and genetic drift and subsequent detrimental effects to genetic diversity (Cabrera-Toledo *et al.*, 2008).

CONSERVATION RECOMMENDATIONS AND CONCLUSIONS

Zamia lucayana is currently categorized as Endangered on the IUCN Red List (IUCN, 2011). We propose that it should be recategorized as Critically Endangered based on criteria B1ab(i–v) + 2ab(i–v). This recommendation has been submitted to the IUCN Cycad Specialist Group. This evaluation is based on the species' highly restricted extent of occurrence (1 km^2) and area of occupancy (0.06 km^2). In addition, *Z. lucayana* requires a unique habitat that is already under residential development.

Zamia lucayana is one of three single-island endemic plants on Long Island (Taylor, 1921, Correll & Correll, 1982). The others are *Euphorbia longinsulicola* and *Matelea correllii.* The former has a relatively narrow distribution range (Hill, 1976) and partially co-occurs with *Z. lucayana*. Because of the few endemics restricted to Long Island, the protection of *Z. lucayana* and its habitat is clearly a major conservation priority for this island.

We recommend both ex situ and in situ conservation for *Zamia lucayana*. The occurrence of the species on private land could impede establishment of an effective management plan. Long-term conservation will therefore depend on whether the Bahamas National Trust and/or the Ministry of the Environment of the Bahamas can purchase the land where this species occurs. Until the required funds are available we recommend the following actions: (1) limit sand mining, (2) plan residential development on central eastern regions of the island that is compatible with the area where *Z. lucayana* occurs, (3) establish national ex situ conservation collections in the Bahamas, with duplicates in other sites, (4) increase conservation awareness

for the species through environmental education programmes, and (5) develop voluntary agreements with landowners to limit development on their property.

During our visit we collected a total of 910 seeds for ex situ conservation from 17 individual plants (Table 1) of *Zamia lucayana*. In addition, we collected total of 75 seedlings and 2,797 seeds from 65 plants of *Z. angustifolia* and *Z. integrifolia*. The location of each plant was determined and sent to the Ministry of the Environment of the Bahamas. The collections were based on the protocols developed at the Montgomery Botanical Center (Walters, 1999, Namoff *et al.*, 2010). In each population we collected seed from five plants and a minimum of 50 seeds per individual. Progeny from each plant is accessioned separately for subsequent ex situ conservation planting. This germplasm has been distributed to the Bahamas National Trust, Fairchild Tropical Botanic Garden, Montgomery Botanical Center, the National Germplasm Repository of USDA-ARS at Miami, the Jardín Botánico Francisco Javier Clavijero (Xalapa, Mexico), and Fairylake Botanical Garden (Shenzhen, China). Among these institutions, the Montgomery Botanical Center maintains one of the most comprehensive living collections of Cycadales worldwide, especially of *Zamia* (Calonje, Husby & Griffith, 2009). The seeds we collected had an 85% germination rate, suggesting that an ex situ conservation program will be feasible.

ACKNOWLEDGMENTS

This study was supported by the Mohamed Bin Zayed Species Conservation Fund (project number 0925331) to JFO, MC, AM, and Tamica Rahming (Bahamas National Trust). Matching funds were provided by Montgomery Botanical Center. This project was also supported by the National Science Foundation (award number 1050340) and by the Christiane Tyson Research Fellowship. We thank C. Adair, R. Adams, C. Calonje, S. Gilmer and L. Johnson for their technical assistance. Ethan Freid provided information pertinent to Long Island endemics. We thank the Ministry of the Environment of the Bahamas for Research and Collection Permits and the Bahamas Department of Agriculture for CITES, export and phytosanitary permits. This is contribution 218 of the Tropical Biology Programme of Florida International University.

LIST OF REFERENCES

Acevedo-Rodríguez, P, Strong, MT. 2012. Catalogue of seed plants of the West Indies. *Smithsonian Contributions to Botany* **98:** 1–1192.

Albury, P 1975. *The Story of the Bahamas* New York, USA: St. Martin's Press.

Anonymous 2002. *Bahamas Environmental Handbook* Nassau, The Bahamas: Bahamas Environment, Science and Technology Commission.

Britton, NL. 1907. Contributions to the flora of the Bahama islands. *Bulletin of the New York Botanical Garden* **5:** 311.

Britton, NL. 1909. Contributions to the flora of the Bahamas Islands IV. *Bulletin of the New York Botanical Garden* **5:** 311–318.

Britton, NL, Millspaugh, CF 1920. *The Bahama flora* New York: The Authors.

Cabrera-Toledo, D, González-Astorga, J, Nicolalde-Morejón, F, Vergara-Silva, F, Vovides, AP. 2010. Allozyme diversity levels in two congeneric Dioon spp. (Zamiaceae, Cycadales) with contrasting rarities. *Plant Systematics and Evolution* **290:** 115-125.

Cabrera-Toledo, D, Gonzalez-Astorga, J, Vovides, AP. 2008. Heterozygote excess in ancient populations of the critically endangered Dioon caputoi (Zamiaceae, Cycadales) from central Mexico. *Botanical Journal of the Linnean Society* **158:** 436-447.

Calonje, M, Husby, C, Griffith, P. 2009. The cycad collection at Montgomery Botanical Center. *Public Garden* **24:** 25–27.

CITES. 2011. The CITES Appendices. (http://www.cites.org/eng/app/index.php) Accessed 25 September 2011.

Cornuet, JM, Luikart, G. 1996. Description and Power Analysis of Two Tests for Detecting Recent Population Bottlenecks From Allele Frequency Data. *Genetics* **144:** 2001-2014.

Correll, DS, Correll, HB. 1982. Zamia. *Flora of the Bahama Archipelago (Including the Turks and Caicos Islands).* Gantner Verlag.

Crawford, NG. 2010. smogd: software for the measurement of genetic diversity. *Molecular Ecology Resources* **10:** 556-557.

Di Rienzo, A, Peterson, AC, Garza, JC, Valdes, AM, Slatkin, M, Freimer, NB. 1994. Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences* **91:** 3166-3170.

Earl, DA, vonHoldt, BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4:** 359-361.

Eckenwalder, J. 1980. Taxonomy of the West Indian cycads. *Journal of the Arnold Arboretum* **61:** 701-722.

Edwards, KJ, Barker, JHA, Daly, A, Jones, C, Karp, A. 1996. Microsatellite Libraries Enriched for Several Microsatellite Sequences in Plants. *BioTechniques* **20:** 758-760.

Evanno, G, Regnaut, S, Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14:** 2611-2620.

Excoffier, L, Lischer, HEL. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10:** 564-567.

Geiger, JH, Meerow, AW, Lewis, C, Oviedo, R, Francisco-Ortega, J. 2014. Genetic diversity and conservation of Ipomoea microdactyla (Convolvulaceae): an endemic vine from the Bahamas, Cuba, and southeastern Florida. *Plant Species Biology* **29:** 2-15.

González-Géigel, L. 2003. Zamiaceae. *Flora de la República de Cuba.* Liechtenstein: A. R. Gantner Verlag KG.

Guo, SW, Thompson, EA. 1992. Performing the Exact Test of Hardy-Weinberg Proportion for Multiple Alleles. *Biometrics* **48:** 361-372.

Hill, SR. 1974. Range extensions and new records for the Bahama flora. *Rhodora* **76:** 471-477.

Hill, SR. 1976. A new species of Euphorbia (Section Chamaesyce) from the Bahamas. *SIDA, Contributions to Botany* **6:** 313-316.

Höglund, J. 2009. Evolutionary conservation genetics.

IUCN. 2011. The IUCN red list of threatened species v. 2011.2 (http://www.iucnredlist.org/) Accessed 17 January 2012.

Jakobsson, M, Rosenberg, NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23:** 1801-1806.

Jost, L. 2008. GST and its relatives do not measure differentiation. *Molecular Ecology* **17:** 4015-4026.

Kimura, M, Crow, JF. 1964. THE NUMBER OF ALLELES THAT CAN BE MAINTAINED IN A FINITE POPULATION. *Genetics* **49:** 725-738.

Kumar, S, Skjæveland, Å, Orr, RJ, Enger, P, Ruden, T, Mevik, B-H, Burki, F, Botnen, A, Shalchian-Tabrizi, K. 2009. AIR: A batch-oriented web program package for construction of supermatrices ready for phylogenomic analyses. *BMC Bioinformatics* **10:** 357.

Langella, O. 2010. Populations 1.2.31 (http://bioinformatics.org/populations/) Accessed 17 January 2012.

Mantel, N. 1967. The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Research* **27:** 209-220.

Maunder, M, Leiva, A, Santiago-Valentín, E, Stevenson, DW, Acevedo-Rodríguez, P, Meerow, AW, Mejía, M, Clubbe, C, Francisco-Ortega, J. 2008. Plant Conservation in the Caribbean Island Biodiversity Hotspot. *The Botanical Review* **74:** 197-207.

Meerow, AW, Francisco-Ortega, J, Calonje, M, Griffith, MP, Ayala-Silva, T, Stevenson, DW, Nakamura, K. 2012. *Zamia* (Cycadales: Zamiaceae) on Puerto Rico: Asymmetric genetic differentiation and the hypothesis of multiple introductions. *American Journal of Botany* **99:** 1828-1839.

Meerow, AW, Nakamura, K. 2007. Ten microsatellite loci from Zamia integrifolia (Zamiaceae). *Molecular Ecology Notes* **7:** 824-826.

Meerow, AWS, D. W.;Moynihan, J.;Francisco-Ortega, J. 2007. Unlocking the coontie conundrum: the potential of microsatellite DNA studies in the Caribbean *Zamia pumila* complex (Zamiaceae). *Memoirs of the New York Botanical Garden* **98:** 484-518.

Namoff, S, Husby, CE, Francisco-Ortega, J, Noblick, LR, Lewis, CE, Griffith, MP. 2010. How well does a botanical garden collection of a rare palm capture the genetic variation in a wild population? *Biological Conservation* **143:** 1110-1117.

Namoff, S, Veloz, A, Jiménez, F, Rodríguez-Peña, RA, Peguero, B, Lewis, C, Moynihan, J, Abdo, M, Maunder, M, Wettberg, EV, Meerow, AW, Griffith, MP, Francisco-Ortega, J. 2011. Sweet Drinks Are Made of This: Conservation Genetics of an Endemic Palm Species from the Dominican Republic. *Journal of Heredity* **102:** 1-10.

Nei, M, Tajima, F, Tateno, Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19:** 153-170.

Norstog, KJ, Nichols, TJ 1997. *The biology of the cycads* Ithaca, New York: Cornell University Press.

Osborne, R. 1995. The world cycad census and a proposed revision of the threatened species status for cycad taxa. *Biological Conservation* **71:** 1-12.

Osborne, R, Calonje, M, Hill, KD, Stanberg, L, Stevenson, DW. 2012. The world list of cycads. *Mem. New York Bot. Gard.* **106:** 480-510.

Peakall, ROD, Smouse, PE. 2006. genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6:** 288-295.

Peterson, PM, McCracken, CL. 2005. Genetic consequences of reduced diversity: heterozygosity loss, inbreeding depression, and effective population size. In: Krupnick, GA, Kress, WJ eds. *Plant Conservation. A Natural History Approach* Chicago, USA: The University of Chicago Press.

Pinares, AG-A, J.;Vovides, A. P.;Lazcano, J.;Vendrame, W. A. 2009. Genetic diversity of the endangered endemic Microcycas calocoma (Miq.) A. DC (Zamiaceae, Cycadales): Implications for conservation. *Biochemical Systematics and Ecology* **37:** 385-394.

Pritchard, JK, Stephens, M, Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155:** 945-959.

Rosenberg, NA. 2004. distruct: a program for the graphical display of population structure. *Molecular Ecology Notes* **4:** 137-138.

Rousset, F, Raymond, M. 1995. Testing heterozygote excess and deficiency. *Genetics* **140:** 1413-1419.

Slatkin, M, Excoffier, L. 1996. Testing for linkage disequilibrium in genotypic data using the Expectation-Maximization algorithm. *Heredity (Edinb)* **76 (Pt 4):** 377-383.

Smouse, PE, Long, JC. 1992. Matrix correlation analysis in anthropology and genetics. *American Journal of Physical Anthropology* **35:** 187-213.

Smouse, PE, Long, JC, Sokal, RR. 1986. Multiple Regression and Correlation Extensions of the Mantel Test of Matrix Correspondence. *Systematic Biology* **35:** 627-632.

Stevenson, DW. 1987. Again the West Indian zamias. *Fairchild Tropical Garden Bulletin* **42:** 23-27

Taylor, N. 1921. Endemism in the Bahama Flora. *Annals of Botany* **35:** 523-532.

Thiers, B. 2011. Index Herbariorum: A global directory of public herbaria and associated staff (Http://sweetgum.nybg.org/ih/) Accessed 10 September 2011.

Walters, TW. 1999. Off-site cycad preservation: II. How to develop off-site collections. *The Cycad Newsletter* **22:** 9-12.

Walters, TW, Decker Walters, DS. 1991. Patterns of allozyme diversity in the West Indies cycad *Zamia pumila* (Zamiaceae). *American Journal of Botany* **78:** 436-445.

Ware, D, Jaiswal, P, Ni, J, Pan, X, Chang, K, Clark, K, Teytelman, L, Schmidt, S, Zhao, W, Cartinhour, S, McCouch, S, Stein, L. 2002. Gramene: a resource for comparative grass genomics. *Nucleic Acids Research* **30:** 103-105.

Yang, Y, Li, Y, Li, LF, Ge, XJ, Gong, X. 2008. Isolation and characterization of microsatellite markers for Cycas debaoensis Y. C. Zhong et C. J. Chen (Cycadaceae). *Molecular Ecology Resources* **8:** 913-915.

Table 1. The six islands of the Bahamas archipelago (Fig. 1) on which cycads of the genus *Zamia* occur, with the species present, and the number of populations studied, seeds collected, and individuals sampled for DNA studies.

Island	Species	No. of populations studied	No. of seeds collected (number of mother plants sampled for seeds)/number of seedlings collected	No. of individuals sampled for DNA studies
Abaco*	*Z. integrifolia*	5	798 (16)/0	76
Andros	*Z. integrifolia*	6	497 (8)/0	158
Eleuthera	*Z. angustifolia*	2	400 (11)/40	37
Eleuthera	*Z. integrifolia*	2	236 (8)/18	46
Grand Bahama	*Z. integrifolia*	1	196 (8)/0	45
Long Island	*Z. lucayana*	3	910 (16)/0	114
New Providence	*Z. integrifolia*	5	670 (14)/17	125
Total		24	3,707 (81)/75	602

*Includes the Tilloo Cay site

Table 2. Main demographic features and descriptive genetic diversity statistics for the three populations of *Zamia lucayana*

	Hamilton's	Buckley's	Petty's
Population code	ZBLI1	ZBLI2	ZBLI3
Number of adults	400	250	300
Number of juveniles	267	643	828
Number of females	120	53	144
Number of males	160	125	120
Number of sterile adults (6 or more leaflet pairs)	120	72	36
Approximate area of occupancy (km^2)	0.45	0.13	0.06
Number of individuals sampled for DNA studies	46	43	33
Average number of alleles per locus	5.3	4.5	4.3
Number of private alleles	15	15	7
Observed heterozygosity	0.519	0. 478	0.454
Expected heterozygosity	0.501	0.486	0.483
Inbreeding coefficient	0.040	0.090	0.053
Percentage of paired loci in linkage disequilibrium	5	0	1

Table 3. Pairwise D_{est} values (above diagonal) and F_{ST} values (below diagonal) for the three populations of *Zamia lucayana* (Table 2).

	ZBLI1	ZBLI2	ZBLI3
ZBLI1		0.051	0.045
ZBLI2	0.075		0.040
ZBL13	0.060	0.062	

Table 4. Estimates of the number of migrants per generation and significance (P) of the analysis of molecular variation (AMOVA) between the three populations of *Zamia lucayana*.

Population comparison	No. of migrants	P
ZBLI1–ZBLI2	3.074	<0.001
ZBLI1–ZBLI3	3.936	<0.001
ZBLI2–ZBLI3	3.778	<0.001

Table 5. Analysis of molecular variance for the three populations of *Zamia lucayana*. The results indicate that the majority of genetic variation is found within, rather than among, populations.

Source	df	SS	MS	Est. Var.	%
Among populations	2	51.270	25.635	0.273	7
Within populations	239	911.739	3.815	3.815	93
Total	241	963.008	4.088	100	

Fig 1. Distribution of the three species of *Zamia* in the Bahamas Islands. Each point represents a population included in our field studies. Population symbols overlap on Andros, Eleuthera, and New Providence. Plants from Tilloo Cay, Abaco, were identified as *Zamia* cf. *integrifolia* for this study. See Table 1 for number of populations and species studied on each island. The inset indicates the location of the Bahamas in the Caribbean.

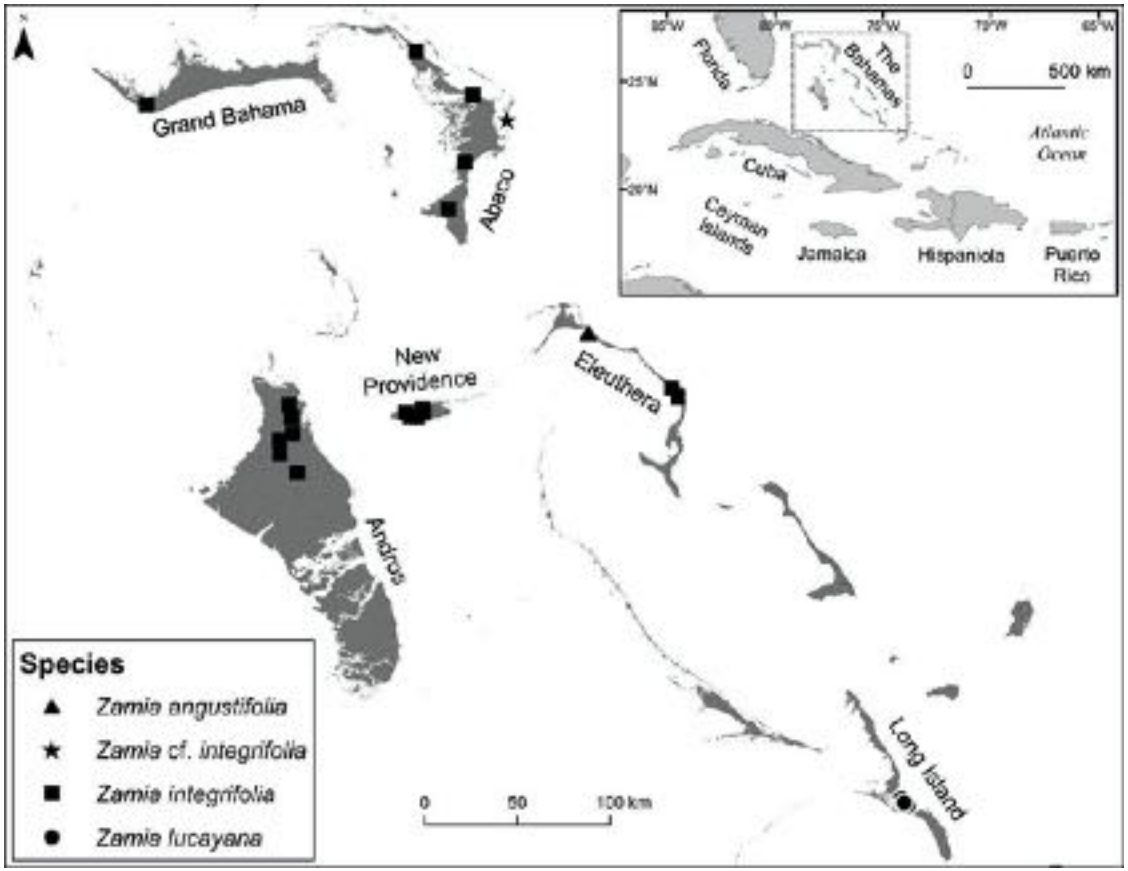


Fig 2. *Zamia lucayana*: (A) male cones, (B) female cone, (C) hermit crab *Coenobita clypeatus* feeding on the fleshy seed coat, (D) adult male individual.

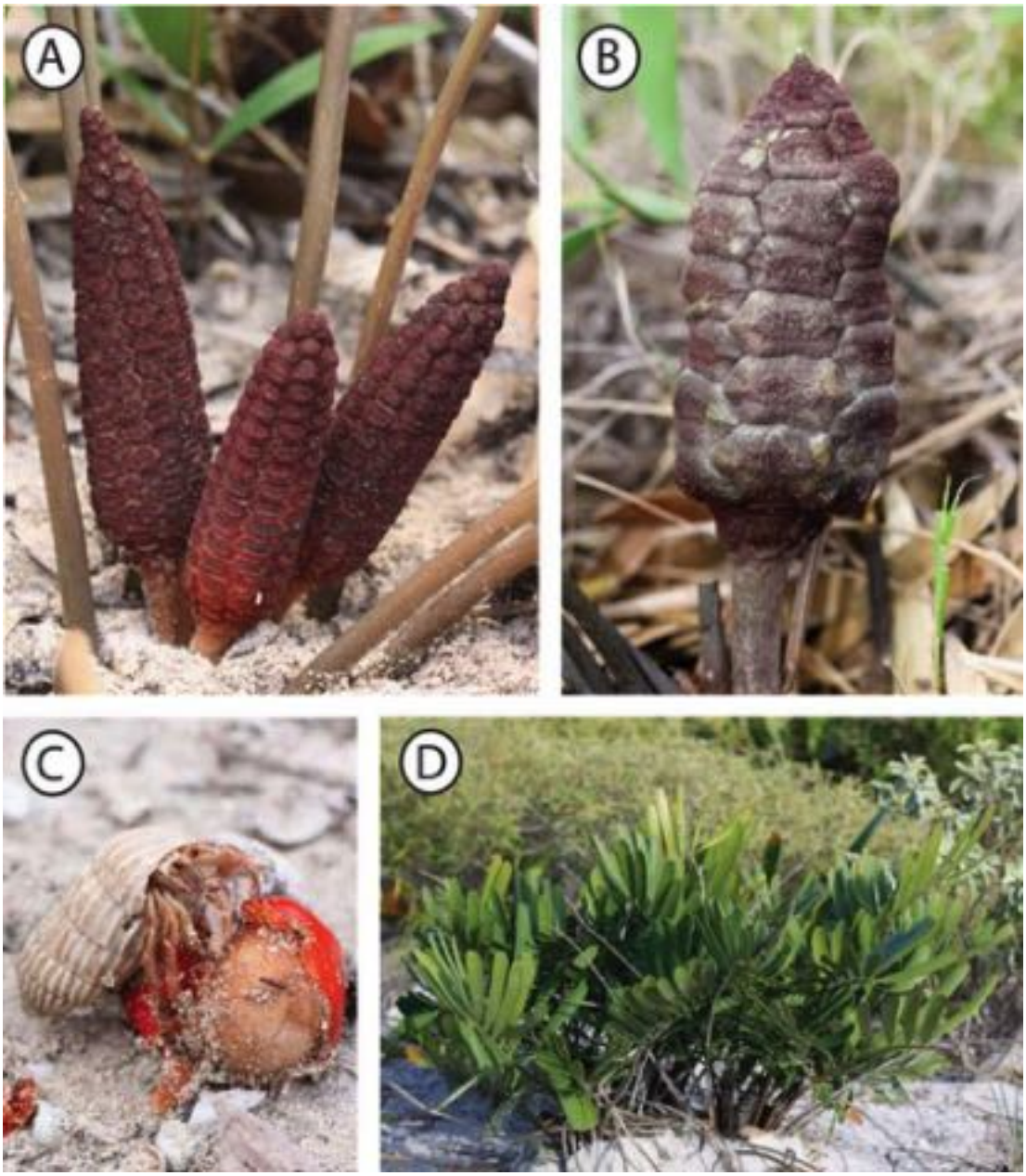


Fig. 3. Distribution of *Zamia lucayana* on Long Island, showing the three major populations included in the conservation genetic study (Buckley's, ZBLI2; Hamilton's, ZBLI1; Petty's, ZBLI3) and two further localities (Galloway Landing, where 10 adult plants were observed, and Mangrove Bush, where 20 adult plants were observed). Each dot represents a site where individual leaflets were sampled for genetic studies (Table 2). No material was sampled at Galloway Landing or Mangrove Bush.

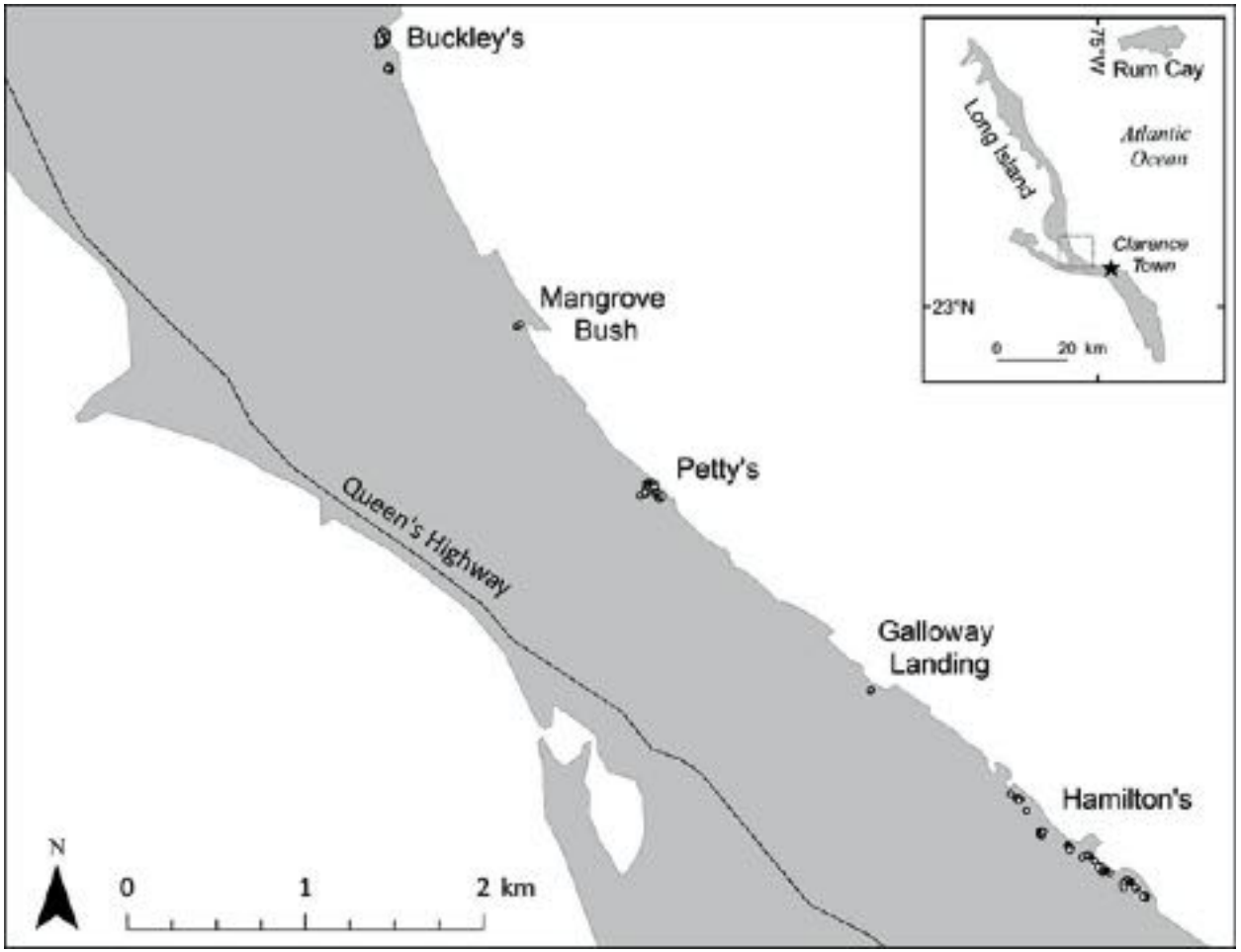
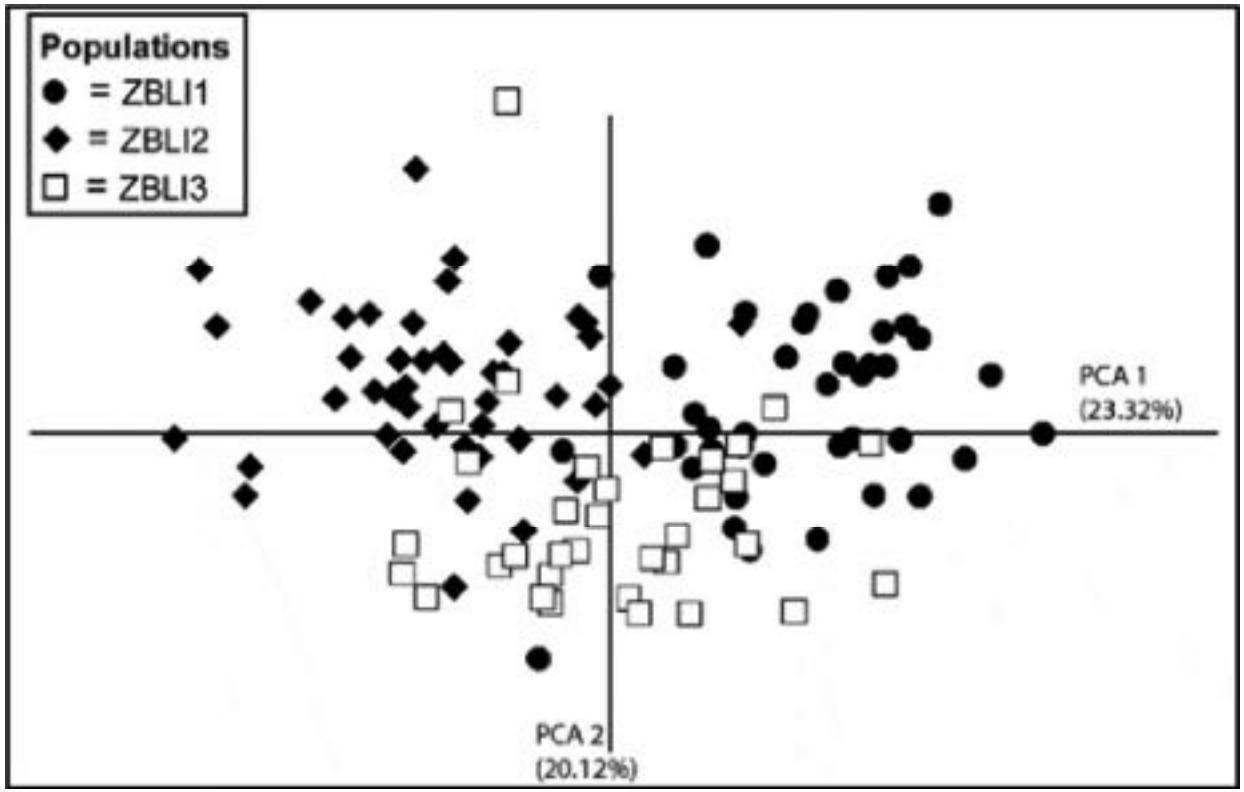


Fig. 4. Graphical representation of genetic structure across three populations of *Zamia lucayana* using Bayesian clustering of microsatellite frequency data. K = 2 was found to be the optimal partition of genetic variation. The individuals of ZBLI1 and ZBLI3 are predominantly assigned to the same cluster (light grey), whereas ZBLI2 is primarily assigned to the second (dark grey). Admixture between the two clusters is indicated by varying proportions of the opposing colour in individuals of each population. ZBLI3 shows a greater proportion of admixture with ZBLI2 than does ZBLI1.

Fig. 5. The first two axes (PCA1, PCA2, with the percentage variation explained by each) of a principal coordinate analysis of DNA microsatellite data for the three main populations of *Zamia lucayana* (Table 2, Fig. 2). Each point represents a single individual.



CONCLUSIONS

In chapter I, a species-comprehensive, well-resolved hypothesis of phylogenetic relationships within the genus *Zamia* was presented. The results suggest a crown age of ~10 Ma for the genus and an origin within the Mesoamerican or Caribbean region with eventual southward migration into the Central American Isthmus and South America. The diversification history of the genus appears relatively stable and characterized by gradual species accumulation and low extinction rates. Most extant *Zamia* species appear to have originated during the Pleistocene, suggesting an effect of climatic oscillations during this epoch on the diversification of the genus. The highest diversification rates were found in clades comprised of Caribbean and South American species. The biogeographic analysis suggests a Caribbean or Mesoamerican origin for *Zamia* with subsequent dispersal to the Central American Isthmus and South America, where the genus reaches its maximum species and morphological diversity. We recovered a strong congruence between geographic distribution and phylogenetic relationships, a pattern consistent with the limited dispersal and pollen-transfer abilities found in cycads. Homoplasy was pervasive within macromorphological characters normally considered informative for the genus, suggesting that future classification work within the genus must rely mostly on genetic characters or explore a wider variety of potentially diagnostic characters, including additional morphological, anatomical, and reproductive characters. This study presents the most comprehensive interpretation of the evolutionary history of *Zamia* to date and should serve as a strong phylogenetic framework for subsequent research, including studies that may help increase our understanding of the mechanistic bases for the diversification of the genus. Future phylogenetic work, particularly with new high-throughput sequencing technologies, should help resolve some of the ambiguities that remain in obtaining a fully resolved, well-supported phylogeny of the genus.

In chapter II I examined the population genetics of *Zamia decumbens,* a Belizean endemic species belonging to a clade of rainforest dwelling species that appears to have primarily diversified during the Pleistocene. Its demographic history appears closely linked to the karstification of the limestone bedrock on which it occurs as well as with the formation of caves necessary to form the collapse features (i.e. sinkholes) in which the largest, most reproductively active known populations occur. The two sinkhole

populations studied here are the largest known population of this species, and our results suggest that they were the first to be established. We found that the genetic variation of the studied populations was structured neither geographically nor ecologically, but rather seemed to reflect the demographic history of the populations and their genetic connectivity. The latter appears to be largely dependent on the Cave population which serves as a hub genetically linking all other populations together, primarily as a source of first-generation migrants to other populations. Future research should focus on identifying the seed dispersal agents for the seeds of this species in hopes of gaining a better understanding of the reasons behind the unusual genetic variation patterns found in this species.

Despite occurring in small isolated populations within the Maya Mountains and considered Critically Endangered, most known populations of *Z. decumbens* occur in protected areas, and the populations surveyed appear reasonably healthy. All populations were in Hardy-Weinberg equilibrium, moderately heterozygous, and did not exhibit decreased heterozygosity, but signatures for recent bottleneck events were recovered for the doline populations. Anecdotal evidence of commercial extraction of mature plants from the Hilltop population appears to be corroborated by the high average pairwise relatedness and inbreeding values recovered for this population in the COANCESTRY analysis. Ex-situ conservation efforts have successfully captured a majority of the allelic diversity of the two sinkhole populations. Future efforts should focus on representing other populations in ex-situ collections and in the propagation of plants in existing collections to make seeds available to the horticultural industry and therefore minimize the demand for wild-collected germplasm. The establishment of local inter-situ sustainable management nurseries for propagating cycads should also be explored, as this model has been previously implemented in Mexico with some success in protecting cycad habitats, discouraging illegal collecting, and benefiting local villages.

In chapter III, I evaluated the diversity of the three species of *Zamia* ocurring in the Bahamas archipelago by surveying twenty-four populations on six different islands, and evaluating their conservation status, habitat preferences, and geographic distribution. *Zamia angustifolia* is endemic to the coastal sand dunes of Eleuthera, and is the species of the highest conservation concern due to its restricted distribution, small total number of plants, and extensive threats from real estate development within its habitat. *Zamia*

integrifolia is the most widespread species, occurring in coastal thicket on Eleuthera, in pinelands and coastal scrub in Abaco, and in pinelands and hardwood coppice in Grand Bahama, New Providence, and Andros. Its distribution is restricted on Eleuthera and Grand Bahama and, although threatened by urban development in New Providence, it is relatively common on Abaco and Andros. *Zamia lucayana* is endemic to coastal sand dunes on Long Island. It is comprised of only three populations within a narrow strip of land of c. 1 km^2; and is very threatened due to its very restricted distribution, small number of populations, and the threat of residential development and sand mining in its coastal habitat. We propose a reassignment of its current conservation status from Endangered to Critically Endangered. Assessment of the genetic structure of *Z. lucayana* indicated that the three known populations are genetically similar and should be considered a single management unit. However, the high number of private alleles suggests that genetic drift, indicative of recent fragmentation, is progressing and will likely continue. A combination of both ex situ and in situ conservation activities were recommended for *Zamia lucayana*, including (1) limiting sand mining, (2) planning residential development on central eastern regions of the island that do not affect the existing populations, (3) establishing national ex situ conservation collections in the Bahamas with duplicates in other sites, (4) increasing conservation awareness for the species through environmental education programs, and (5) developing voluntary agreements with landowners to limit development on their property.

The three chapters together provide novel insights into the evolutionary genetics of the genus *Zamia*. The first chapter by presenting a broad scale phylogenetic framework for the genus as a whole, the following two chapters by providing finer-scale insights at the population level for two species (*Z. decumbens* and *Z. lucayana*) belonging to separate clades and ocurring in very distinct habitats.

VITA

MICHAEL CALONJE BAZAR

Born, Cali, Colombia

2000	B.A., Cinema San Francisco State University San Francisco, California
2005	M.S., Environmental Systems Humboldt State University Arcata, California
2004-2007	Environmental Consultant Rabe Consultting Klamath Falls, OR
2019	Ph.D. Candidate, Biology Florida International University Miami, Florida

PUBLICATIONS

Segalla, R. and M. Calonje. 2019. *Zamia brasiliensis,* a new species of *Zamia* (Zamiaceae, Cycadales) from Mato Grosso and Rondônia, Brazil. Phytotaxa 404(1): 1-11. DOI: 10.11646/phytotaxa.404.1.1

Calonje, M., A.W. Meerow, M.P. Griffith, D. Salas-Leiva, A.P. Vovides, M. Coiro, and J. Francisco-Ortega. 2019. A time-calibrated species tree phylogeny of the New World cycad genus *Zamia* L. (Zamiaceae, Cycadales). International Journal of Plant Sciences 180(4): 286-314. DOI: 10.1086/702642

Calonje, M., C. López-Gallego, and J. Castro. 2018. *Zamia paucifoliolata*, a new species of *Zamia* (Zamiaceae, Cycadales) from Valle del Cauca, Colombia. Phytotaxa 385 (2): 85-93. DOI: http://dx.doi.org/10.11646/phytotaxa.385.2.4

Erdei, B., M. Calonje, A. Hendy, and N. Espinosa. 2018. A review of the Cenozoic fossil record of the genus *Zamia* L. (Zamiaceae, Cycadales) with recognition of a new species from the late Eocene of Panama – evolution and biogeographic inferences. Bulletin of Geosciences 93(2): 185-204.

Stevenson, D.W., L. Stanberg and M. Calonje. 2018. The World List of Cycads. Memoirs of the New York Botanical Garden 117: 540-576

Meerow, A.W, Salas-Leiva, D.E., M. Calonje, J. Francisco-Ortega, M.P. Griffith, K. Nakamura, F. Jiménez-Rodríguez, J. Lawrus, and A. Oberli. 2018. Contrasting Demographic History and Population Structure of Zamia (Cycadales: Zamiaceae) on Six Islands of the Greater Antilles Suggests a Model for Population Diversification in the Caribbean Clade of the Genus. International Journal of Plant Sciences. DOI 10.1086/700104.

Calonje, M. and A. N. Sennikov. 2017. In the process of saving plant names from oblivion: The revised nomenclature of *Ceratozamia fuscoviridis* (Zamiaceae). Taxon 66(1):158-164.

Calonje, M., G. Morales, C. Lopez-Gallego, and F. J. Roldan. 2015. A taxonomic revision of *Zamia montana* and *Zamia oligodonta*, with notes on their conservation status. Phytotaxa 192: 279-289.

Calonje, M., A. W. Meerow, L. Knowles, D. Knowles, M. P. Griffith, K. Nakamura, J. Francisco-Ortega. 2013. Cycad biodiversity in the Bahamas Archipelago and conservation genetics of the Critically Endangered *Zamia lucayana* (Zamiaceae) Oryx 47: 190-198.

Salas-Leiva, D., A. W. Meerow, M. Calonje, M. P. Griffith, J. Francisco-Ortega, D. W. Stevenson, K. Nakamura, C. E. Lewis, S. Namoff. 2013. Phylogeny of the cycads based on multiple single copy nuclear genes: congruence of concatenation and species tree inference methods. Annals of Botany. doi: 10.1093/aob/mct192.

Calonje, M., H. E. Esquivel, G. Morales, Y. A. Mora-Lizcano, and D. Stevenson. 2012. A new arborescent species of *Zamia* (Cycadales, Zamiaceae) from the Department of Huila, Eastern Cordillera of Colombia. Caldasia 34: 283-290.

Calonje, M. A., A. S. Taylor Blake, D. Wm. Stevenson, G. Holzman, and Y. A. Ramos. 2012. *Zamia lindleyi*, a misunderstood species from the highlands of Western Panama. Memoirs of the New York Botanical Garden 106: 419-437.

Calonje, M., H. E. Esquivel, D. Stevenson, C. Calonje and D. Pava. 2011.A new arborescent species of Zamia from the Central Cordillera of Tolima, Colombia (Cycadales, Zamiaceae), with comments on the *Z. poeppigiana* species complex. Brittonia 63: 442-451.

Calonje, M., A. W. Meerow, and D. W. Stevenson. 2010. The typification of *Zamia erosa* and the priority of that name over *Z. amblyphyllidia*. Taxon 59:637-642.

Calonje, M., D. Stevenson, C. Calonje, Y. A. Ramos, and A. Lindstrom. 2010. A new species of *Zamia* from Chocó, Colombia (Cycadales, Zamiaceae). Brittonia 62: 80-85.